

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 19, 2004, 18:54:54 ; Search time 7554 Seconds
(without alignments)
11905.382 Million cell updates/sec

Title: US-10-619-685-1
Perfect score: 2468
Sequence: 1 aaaaaattttcaatcac.....aaaaaaaaaaaaaaaaaaaaa 2468

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gest1:*
9: gb_gest2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	555.4	22.5	565	CF323765	CF323765 HDN--04-L
2	551	22.3	753	AG021802	AG021802 Oryza sat
3	542.4	22.0	564	CF323815	CF323815 HDN--04-N
4	494.8	20.0	706	CL833850	CL833850 OR CBA005
5	463	18.8	646	CA498690	CA498690 WHE3246_E
6	446.2	18.1	604	CA031067	CA031067 HX08017i
7	446	18.1	627	EU037770	EU037770 946142A06
8	429.2	17.4	1076	CA254411	CA254411 SCBFL411
9	411.8	16.7	667	CA298359	CA298359 SCRFPL803
10	410.4	16.6	617	BG487823	BG487823 FM1_73_H0
11	404.2	16.4	807	CK193855	CK193855 FGAS00227
12	402.4	16.3	548	BG604893	BG604893 WHE0944_F
13	398.4	16.1	1986	CNS0A2G0	BX827777 Arabidops
14	395.6	16.0	2195	CNS09YQA	BX831110 Arabidops
15	394	16.0	699	AL506716	AL506716 AL506716
16	391.8	15.9	675	CA235572	CA235572 SCACFL501
17	391.2	15.9	450	CF330979	CF330979 NACL--06-
18	385	15.6	526	CA251300	CA251300 SCOSFL112
19	376	15.2	795	CD979653	CD979653 QAG5D03.X
20	375.8	15.2	594	CA232369	CA232369 SCRFPL306
21	375	15.2	545	AV833315	AV833315 AV833315
22	365.4	14.8	641	AV833315	AV833315 AV833315
23	359.8	14.5	494	BJ224222	BJ224222 BJ224222
24	356.8	14.5	719	AY108867	AY108867 Zea mays

25	350.2	14.2	592	2	BF484964	BF484964 WHE2333.F
26	349.4	14.2	635	6	CF038174	CF038174 QCH17F09.
27	347.4	14.1	611	2	BE511523	BE511523 946061D08
28	341.2	13.8	510	4	BG465723	BG465723 RH122_47
29	340.8	13.8	507	6	CA831022	CA831022 1117014E0
30	322.6	13.1	517	6	CA249488	CA249488 SCRFPL111
31	320.4	13.0	504	6	CD979546	CD979546 QAG4305.X
32	319.4	12.9	744	4	BG646008	BG646008 ESTS07627
33	314.2	12.7	758	5	BQ990215	BQ990215 QGF19M11.
34	314	12.7	569	4	BJ478725	BJ478725 BJ478725
35	301	12.2	1583	3	AY104938	AY104938 Zea mays
36	297.6	12.1	400	2	BF484550	BF484550 WHE2324.H
37	297	12.0	562	4	BG158382	BG158382 FM1_59_F1
38	291.8	11.8	510	4	BM097574	BM097574 EBem04_SQ
39	287.8	11.7	584	1	AJ500791	AJ500791 AJ500791
40	283.2	11.5	412	4	BG241888	BG241888 RH122_51
41	282.2	11.4	802	7	CF922842	CF922842 gmrhrw24
42	282	11.4	707	5	BU004982	BU004982 QGG6M05.Y
43	280.6	11.4	550	5	BQ608748	BQ608748 BRY_4659
44	272.8	11.1	533	5	BQ620280	BQ620280 taLr1173G
45	266	10.8	939	8	BZ795282	BZ795282 PUFAN64TD

RESULT 1

CF323765

LOCUS

DEFINITION

CF323765

HDN--04-L11.g1 OSHDAC1-overexpressing transgenic rice lambda phage

CDNA library II (HDN) Oryza sativa (japonica cultivar-group) cDNA

clone HDN--04-L11, mRNA sequence.

ACCESSION

CF323765

VERSION

CF323765.1

KEYWORDS

EST.

SOURCE

Oryza sativa (japonica cultivar-group)

ORGANISM

Oryza sativa (japonica cultivar-group)

REFERENCE

1 (bases 1 to 565)

AUTHORS

Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C., Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

TITLE

Large-scale Sequencing Analysis of Rice ESTs

JOURNAL

Unpublished (2003)

COMMENT

Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division of Bioscience and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.com, bhnahm@bio.myongji.ac.kr.

FEATURES

source

Location/Qualifiers

1..565

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nackdong"

/db_xref="taxon:39947"

/clone="HDN--04-L11"

/tissue_type="callus"

/dev_stage="proliferated callus on 2N6 media for 2 weeks"

/lab_host="E.coli SOLR"

/clone_lib="OSHDA1-overexpressing transgenic rice lambda phage CDNA library II (HDN)"

/notes="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2: XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at 5' end with EcoRI and 3' end with XhoI site. mRNA was derived from rice Histone Deacetylase overexpression line."

ORIGIN

Query Match 22.5%; Score 555.4; DB 6; Length 565;

Best Local Similarity 99.8%; Pred. No. 8.9e-132;

Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 755 TTTGAATCTTTGCAATTCCTTATGTGAGGTGCTTGGCATAGATTTCAAGCAACAGATATA 814
 Db 9 TTTGAATCTTTGCAATTCCTTATGTGAGGTGCTTGGCATAGATTTCAAGCAACAGATATA 68

Qy 815 TGAGGTGCACCTAGCTTGGACGAAGCTGAAGGATCAAGAAACCTTGACCAACACTACAAAT 874
 Db 69 TGAGGTGCACCTAGCTTGGACGAAGCTGAAGGATCAAGAAACCTTGACCAACACTACAAAT 128

Qy 875 TGAGAGGCTTGTCTGCTGCCGCAACAGACTCGGTGAATGAAGATCCAAAGATGCAAAA 934
 Db 129 TGAGAGGCTTGTCTGCTGCCGCAACAGACTCGGTGAATGAAGATCCAAAGATGCAAAA 188

Qy 935 GCTTCAAGATTTGCTTCTAGCATGCTCGAGTATGGAATCTCATGGATCTCCACTTGA 994
 Db 189 GCTTCAAGATTTGCTTCTAGCATGCTCGAGTATGGAATCTCATGGATCTCCACTTGA 248

Qy 995 AGAGCAGCAGATGTTTTCAGAAATATAACATGCAATATTCCTGCTTCAGAAACAGAGATAAC 1054
 Db 249 AGAGCAGCAGATGTTTTCAGAAATATAACATGCAATATTCCTGCTTCAGAAACAGAGATAAC 308

Qy 1055 TGAACCAACACCTCTCCACAGATTTCTCTGAATATGCAATCTGAGGTGTTAAGGCT 1114
 Db 309 TGAACCAACACCTCTCCACAGATTTCTCTGAATATGCAATCTGAGGTGTTAAGGCT 368

Qy 1115 TGAACCACTGAAGCAAGTATGATGAAGATCTGTTTAAAGAAAGAGAGACTAGA 1174
 Db 369 TGAACCACTGAAGCAAGTATGATGAAGATCTGTTTAAAGAAAGAGAGACTAGA 428

Qy 1175 AGAGCATAGAAGACGTGCTCATCTTGTGGGAGGAGAGGTTATGAGAGGAGTTTAGCAT 1234
 Db 429 AGAGCATAGAAGACGTGCTCATCTTGTGGGAGGAGAGGTTATGAGAGGAGTTTAGCAT 488

Qy 1235 TGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTGTAACAAATGGAAGCTCA 1294
 Db 489 TGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTGTAACAAATGGAAGCTCA 548

Qy 1295 CATTGCAACAGTGAAG 1311
 Db 549 CATTGCAACAGTGAAG 565

RESULT 2

AG021802

LOCUS

DEFINITION

Oriza sativa (japonica cultivar-group) DNA, clone:NC0266.0.102_1A, 3' flanking sequence of Tos17 insertion in rice strain NC0266, genomic survey sequence.

ACCESSION

AG021802

VERSION

AG021802.1

GI:7679977

KEYWORDS

GSS.

ORGANISM

Oriza sativa (japonica cultivar-group)

Oriza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oriza.

REFERENCE

1

Miyao.A., Tanaka.K., Murata.K., Sawaki.H., Takeda.S., Abe.K., Shinzuka.Y., Onosato.K. and Hirochika.H.

Target site specificity of the Tos17 retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome

Plant Cell 15 (8), 1771-1780 (2003)

JOURNAL

MEDLINE

PUBMED

12897251

REFERENCE

2

(bases 1 to 753)

Direct Submission

Authors

Miyao.A., Tanaka.K. and Hirochika.H.

Submitted (25-OCT-1999)

Akio Miyao, National Institute of Agricultural Sciences, Molecular Genetics; 2-1-2, Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyao@affrc.go.jp, URL:http://tos.nias.affrc.go.jp/, Tel:81-298-38-7020,

FEATURES

source

Location/Qualifiers

1..753

/organism="Oriza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/strain="NC0266"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="NC0266.0.102_1A"

/clone_lib="PCR product directly amplified from rice genomic DNA"

/note="The 3' end of retrotransposon Tos17 was found immediately upstream of this sequence."

misc_feature

1

/note="This insertion point is base 20049 in AP003452."

ORIGIN

Query Match 22.3%; Score 551; DB 9; Length 753;

Best Local Similarity 98.2%; Pred. No. 1.3e-130;

Matches 554; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1891 GTTTGGACATTTGCCGGATTCGCTATCAAGAAGTTGTCTTCAATGCCAGTACTCTACGTG 1950
 Db 1 GTTTGGACATTTGCCGGATTCGCTATCAAGAAGTTGTCTTCAATGCCAGTACTCTACGTG 60

Qy 1951 AGACGGAGACACCTCGTAAACCTTTTGTCTCAGATCACACAGGAAACAGTGTCTGTCGA 2010
 Db 61 AGACGGAGACACCTCGTAAACCTTTTGTCTCAGATCACACAGGAAACAGTGTCTGTCGA 120

Qy 2011 CGCCTGTGCGCCCTTATCAACCAATAACTGAGGATGATGAGAACAGGACTCCGAAGACAT 2070
 Db 121 CGCCTGTGCGCCCTTATCAACCAATAACTGAGGATGATGAGAACAGGACTCCGAAGACAT 180

Qy 2071 TTACAGCAGTGAATCCCAAGACTCCGATGACTGTTACGGCTCCAATGCAGATGCAATGA 2130
 Db 181 TTACAGCAGTGAATCCCAAGACTCCGATGACTGTTACGGCTCCAATGCAGATGCAATGA 240

Qy 2131 CTCCCTCTCTGGCCCAACAAGGTTTCAGCAACTCCAGTTTCCTTGTGTACGACAGCCAG 2190
 Db 241 CTCCCTCTCTGGCCCAACAAGGTTTCAGCAACTCCAGTTTCCTTGTGTACGACAGCCAG 300

Qy 2191 AGGTAACATTTGCAGGAGGACATCGACTACTCTTTGAAGAAAGCGGCTCGCCATCTATC 2250
 Db 301 AGGTAACATTTGCAGGAGGACATCGACTACTCTTTGAAGAAAGCGGCTCGCCATCTATC 360

Qy 2251 TGGCCAGGCAAAATGGTTTAAGTGTGATCAATTTATGTACGTAGTTGAAATCTGACTGCA 2310
 Db 361 TGGCCAGGCAAAATGGTTTAAGTGTGATCAATTTATGTACGTAGTTGAAATCTGACTGCA 420

Qy 2311 TTTTCTTGTGCGTGGCCATTTGCGTATGTTGTTCAACAATAGTCGGCTTTCCAGTAGCAC 2370
 Db 421 TTTTCTTGTGCGTGGCCATTTGCGTATGTTGTTCAACAATAGTCGGCTTTCCAGTAGCAC 480

Qy 2371 TATTCTGATTTACTGCAATTTGTTTAAATGTTTCTACAAACAGTAAACAGCTCTATACA 2430
 Db 481 TATTCTGATTTACTGCAATTTGTTTAAATGTTTCTACAAACAGTAAACAGCTCTATACA 540

Qy 2431 TTAGCTTGTCTCAGTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 2454
 Db 541 TTAGCTTGTCTCAGTAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 564

RESULT 3

CF323815

LOCUS

DEFINITION

HDN--04-N13.g1 OsHDAC1-overexpressing transgenic rice lambda phage cDNA library II (HDN) Oriza sativa (japonica cultivar-group) cDNA clone HDN--04-N13, mRNA sequence.

ACCESSION

CF323815

VERSION

CF323815.1

GI:33795895

KEYWORDS

EST.

SOURCE

Oriza sativa (japonica cultivar-group)

ORGANISM

Oriza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 564)

Kim, J.S., Jun, K.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,

Song, S.I., Kim, J.K., Kim, Y.-K. and Nahm, B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

Contact: Nahm, B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division

of Bioscience and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bhnahm@bio.myongji.ac.kr.

Location/Qualifiers

1..564

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nackdong"

/db_xref="taxon:39947"

/clone="HDN-04-N13"

/tissue_type="callus"

/dev_stage="proliferated callus on 2N6 media for 2 weeks"

/lab_host="E.coli SOLR"

/clone_lib="OshDAC1-overexpressing transgenic rice lambda

phage cDNA library II (HDN)"

/note="Vector: pBluescript SK(+); Site_1: EcoRI; Site_2:

XhoI; cDNA was inserted into lambda Uni-ZAP XR vector at

5' end with EcoRI and 3' end with XhoI site. mRNA was

derived from rice Histone Deacetylase overexpression

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

line."

QY 1294 ACATTGCAACAGTGAA 1309

|||||

Db 549 ACATTGCAACAGTGAA 564

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

RESULT 4

CL833850

LOCUS

DEFINITION

OR CBa0057B05.f OR Cba Oryza rufipogon genomic clone OR_CBa0057B05

5'- genomic survey sequence.

ACCESSION

CL833850

VERSION

CL833850.1

GI:51079460

GSS.

KEYWORDS

ORGANISM

Oryza rufipogon

Oryza rufipogon

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzae; Oryza.

1 (bases 1 to 706)

Kim, H., Yu, Y., Wissotski, M., Yost, D., Stum, D., Rao, K., Luo, M.,

Jetty, R., Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and

Wing, R.

OMAP Project

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

University of Arizona

Forbes Building Room 303, Tucson, AZ 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: http://genome.arizona.edu

PCR Primers

FORWARD: TAA TAC GAC TCA CTA TAG GG

BACKWARD: CAC TCA TTA GSC ACC CCA

Plate: 0057 row: B column: 05

Seq primer: TAA TAC GAC TCA TAG GG

Class: BAC ends.

Location/Qualifiers

1..706

/organism="Oryza rufipogon"

/mol_type="genomic DNA"

/db_xref="taxon:4529"

/clone="OR_CBa0057B05"

/tissue_type="young leaves"

/dev_stage="2 week old seedlings"

/lab_host="DH10B T1 phage resistant"

/clone_lib="OR_CBa"

/note="Vector: pAGIBAC1; Site_1: HindIII; Site_2: HindIII;

drk treated 36 hrs before harvest"

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 607; Conservative

20.0%; Score 494.8; DB 9; Length 706;

86.5%; Pred. No. 4,1e-116; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

Mismatches 0; Mismatches 7; Indels 88; Gaps 2;

```

Db      241 TCACCTGAACATATGCTCAACTCAACCAACAAATATCATGTAAGTGTGGACATG 300
Qy      1903 CCGAATGCTATCAAGAAGTGTCTTCAATGCCAGTACTCTAGTGAGACGAGACAC 1962
Db      301 CCGGATGCTATCAAGAAGTGTCTTCAATGCCAGTACTCTAGTGAGACGAGACAC 360
Qy      1963 CTCGTAACCTTTCTGCTAGATCACACAGGAACAGTGTCTCTGCGAGCTGTGGCC 2022
Db      361 CTCGTAACCTTTCTGCTAGATCACACAGGAACAGTGTCTCTGCGAGCTGTGGCC 420
Qy      2023 CTATCACCATAACACTGAGGATGATGAGAACAGGACTCCGAAAGACATTTACAGCACTGA 2082
Db      421 CTATCACCATAACACTGAGGATGATGAGAACAGGACTCCGAAAGACATTTACAGCACTGA 480
Qy      2083 ATCCCAAGACTCCGATGACTGTTCAGGCTTCAATGAGATGCAATGACTCTCTCTGG 2142
Db      481 ATCCCAAGACTCCGATGACTGTTCAGGCTTCAATGAGATGCAATGACTCTCTCTGG 540
Qy      2143 CCAACAAGGTTTCAGCACTCCAGTTCCTCTTTTACGACAGCCAGGATTAACATGCG 2202
Db      541 CCAACAAGGTTTCAGCACTCCAGTTCCTCTTTTACGACAGCCAGGATTAACATGCG 600
Qy      2203 AGGAGGACATCGAC-TACTCTCTTTGAAGAAAGCGGCTCGCCATCTATCTGGCCAGGCAA 2261
Db      601 AGGAGGACATCGACTTACTCTTTTGAAGAAAGCGGCTCGCCATCTATCTGGCCAGGCAA 660
Qy      2262 ATGGTTTAACTGATCAATTTATGTAGTGTAGTGTAAATCT 2303
Db      661 ATGGTTTAACTGATCAATTTATGTAGTGTAGTGTAAATCT 702

```

```

RESULT 5
CA498690
LOCUS
DEFINITION
WHE3246_E07_I142 Wheat meiotic anther cDNA library Triticum
aestivum cDNA clone WHE3246_E07_I14, mRNA sequence.
ACCESSION
CA498690
VERSION
CA498690.1 GI:24989650
KEYWORDS
EST.
SOURCE
Triticum aestivum (bread wheat)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
REFERENCE
1 (bases 1 to 646)
Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R.,
Pham, J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.
The structure and function of the expressed portion of the wheat
genomes - Meiotic anther cDNA library
Unpublished (2002)
CONTACT: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: oanderson@pw.usda.gov
Sequences have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: T7 primer.
FEATURES
Location/Qualifiers
1..646
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE3246_E07_I14"
/tissue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="Wheat meiotic anther cDNA library"
/note="Vector: pSPOR1; Site_1: SalI; Site_2: NotI; Plants

```

were grown in a glasshouse. Anther meiotic stage was determined by removing anthers from individual primary florets. One anther was sacrificed for microscopic staging, and if determined to be between (and including) meiotic stages pre-meiosis and metaphase I, the remaining two anthers were collected and pooled for library construction. The tissue, total RNA, and poly(A) RNA were prepared, cDNA synthesised, and directionally ligated into pSPOR1 by Tim Sutton in the P Langridge Lab at the Department of Plant Science, University of Adelaide, Waite Campus, Australia. Average insert size 1.5kb. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

ORIGIN

```

Query Match      18.8%; Score 463; DB 6; Length 646;
Best Local Similarity 83.9%; Pred. No. 6.4e-108;
Matches 536; Conservative 0; Mismatches 100; Indels 3; Gaps 1;

Qy      978 ATGATATCTCCACTTGAAGAGCAGCAGATGTTTCCAGATATTAACATCAATATGCTGCT 1037
Db      11 ATGATATCACCAATTTGAAGAGCAGCAGATGTTTCCAGATATTAACATCAATATGCTGCT 70

Qy      1038 TCAGAACAGAGATAACTGAACCAACCAACATCTCTCCACAGATTTCTCGAATATATGTCGAA 1097
Db      71 TCAGAGCATGAATTAACAGAACCCACACACCTCTGATCGACTTCTCAGCTATGTGAA 130

Qy      1098 TCTGAGGTGTTAAGGCTTGAACTGAAGCAAGTAAAGATGAAGATCTTGTGTTTAAAA 1157
Db      131 GCTGAGGTTTTAAGGCTTGAGCAACTGAAGGAAGCAAAATGAAGGAGCTCGTTTGA 190

Qy      1158 AAGAAAGCAGAACTAGAGAGCAGTGAAGAGCTCTCATCTTGTGCGAGGAGGTTAT 1217
Db      191 AAGAAATCGGAGCTGAGAGGACACAGAGAGCTGCGCATCTAATTTGGCGAGGAGGATAT 250

Qy      1218 GCAGAGGAGTTTAGCTTGAAGCTTATTGAAGCTGGAGCTATTGATTCCTCAGTACTT 1277
Db      251 TCAGATCAATTTAATCAATTTAGGCTTATTGAGTGGGAGCTATTGATTCCTGCAATTCGTG 310

Qy      1278 GAACAAATTTAGCTCAGCTTGAAGTGAAGTGAAGAGGAGCTTTAGCCGGAAGATAT 1337
Db      311 GAACAAATTTAGGCTTCACTTTGCAACAGTGAAGATGAAGCTTTTCACTGCAAGGATAT 370

Qy      1338 CTTGAGAAAGTTGAAGATGCAAAATGCTTTGTGAAGGAAGAGCTGGCTGGAAGATTAC 1397
Db      371 CTTGAGAAAGTTGAAGATTTCTGAATGCAATGTGAGGAGGAGAGCTGGCTGGAAGATTAC 430

Qy      1398 AACAAAGATGATATCGTTTACAATGCTGGAGGGGAGCAGATCTTAACACTTAAGAGGGCT 1457
Db      431 AACAAAGATGACAATCGTTACAATGCTGGGAAGGGGAGCAGATCTGACGCTCAAGAGGCT 490

Qy      1458 GAAAGGCTCGTACTTTGGTCAACAGATTTCTCGGAATGGTAGATGTTTGGAGAACAAA 1517
Db      491 GAGAAAGCTCGCATTTTGGTTTAACAGATTTCCGGGAATGGTAGATGTTTGGACACAAA 550

Qy      1518 ATTGCTCATGGAAAAATGAACAGGAGGAGGAGGATTTACATATGATGGTGTAGGCTT 1577
Db      551 ATCATGCTTGGGAGATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 607

Qy      1578 TCGTCAATGCTTGAATATATGTTTCGTTTCGTCAGGAG 1616
Db      608 CTCTCAATGCTGGATGAATATATGATTTGTTTCGCGAGGAG 646

```

```

RESULT 6
CA031067
LOCUS
DEFINITION
HX080171 HX Hordeum vulgare subsp. vulgare cDNA clone HX08017
5-PRIME, mRNA sequence.
ACCESSION
CA031067
VERSION
CA031067.1 GI:24327098
KEYWORDS
Hordeum vulgare subsp. vulgare

```


ORGANISM *Hordeum vulgare* subsp. *vulgare*
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

REFERENCE 1 (bases 1 to 604)
 Zhang, H., Weschke, W., Michalek, W., Stein, N. and Graner, A.
 EST sequencing and analysis in barley (2002)
 Unpublished (2002)
 COMMENT Contact: Stein Nils
 Molecular Markers Group, Department Genbank
 Institute of Plant Genetics and Crop Plant Research (IPK)
 Corrensstr. 3, 06466, Gatersleben, Germany
 Tel: 039482-5522
 Fax: 039482-5595
 Email: stein@ipk-gatersleben.de
 Insert Length: 604 Std Error: 0.00
 Plate: 8 row: 0 column: 17
 Seq primer: M3rev.

FEATURES

source	Location/Qualifiers
	1..604
	/organism="Hordeum vulgare subsp. <i>vulgare</i> "
	/mol_type="mRNA"
	/cultivar="barke"
	/sub_species="vulgare"
	/db_xref="GABI:271738"
	/db_xref="taxon:112509"
	/clone="HX08017"
	/tissue_type="apex"
	/dev_stage="apex (3-5 mm in size)"
	/lab_host="XL10-Gold"
	/clone_lib="HX"
	/note="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); Due to a cloning artefact caused by the kit, in most cases the EcoRI site is NOT present, as well as the EcoRI adapter used for cloning. To excise the insert, restriction sites upstream EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also due to the cloning system used Blue/white selection for recombinants is not 100% reliable."

ORIGIN

Query Match	18.1%;	Score 446.2;	DB 6;	Length 604;
Best Local Similarity	83.7%;	Pred. No. 1.4e-103;		
Matches 505;	Conservative	0;	Mismatches 98;	Indels 0; Gaps 0;

Qy 276 TGGGATGAGTCCGTGAGCCGACACGACGAGGACAGGATGCTGCTGAGACTCGACGAG 335
 Db 2 TGGGACGAGTCCGCGAGCCCGACTCGGTGCGGACACAGACGCTGCTGGAGCTAGACGAG 61

Qy 336 GAGTGCCTGGAGTCTTACAGGCGGAAGGTCCACAGGCGGACCGAGCCGCGCCAGCTG 395
 Db 62 GAGTGTCTCGAGCTCTACAGGAGGAAGGTGCACAGGCGCAACCCGCTGACGGCTCAGCTG 121

Qy 396 CGGAAGGCCATCGCCGAGGCGGACGACAGCTCGCCGGCATCTGCTCAGCCATCGGCGAG 455
 Db 122 CGCAATCCATCGGCGAGCCCGAGCCGAGCTCGCGGTATCTGTCGGCCATAGGCGAG 181

Qy 456 CGCCCGGTGCACTTTAGACAGTCAATCAGAAGCTTCATGCTTAAGAGAGGAGTTGAAT 515
 Db 182 CCACCGGTTTCAATCAGGCAGTCAAAACAGAAATTTTACATGGTTTTAAGGGAGGAGTTGAAT 2411

Qy 516 GCAATTTGTTCCGTTATTTGGAAGAAATGAAAAAGAAAAAGGTTCGAACGATGGAACGAGTTT 575
 Db 242 GCAATTAATCCATATTTGAGGAGATTCGAACTAAAGAGTTTGAAGATGGAACGAGTTC 301

Qy 576 GTTCATGCTCATAGACAGATTTAAGAAAAATTTTCGTCTGAAATAAGGCCAGCCGATTTTGT 635
 Db 302 GTTCGTCTCTTCAGGAGATTAAGAGATTTTCATCCGAAATCAGGCCCTGCAGACTTTGTG 361

Qy 636 CCCTTTAAAGTTCGGTGTGATCAGTCTGACCTGTCTATTAAGAAAGCTTGATGAGTGAACG 695
 Db 362 CCATTTAAACTCTCTGTGATCAGTCTGATTTGTCTATTAAGAAAGTTTGAGGAGCTTAACG 4211

```
QY 1192 CTCTCTTTGTCGCGAGGAGCTTATGACAGAGCTTTAGCATTGAAGCTATTGAAGCTG 1251
Db 121 CTCTCTTTGTCGCGAGGAGGCTATGACAGCTGAATTTTAGCATGAGCTATTGAGGCG 180
QY 1252 GAGCTATTGATCCCTCACCTAGTACTTGAACAAATTTGAAGCTCACATTGCAACAGTGAAG 1311
Db 181 GAGCTTTGATCTCTGCTGCTGCTGGAACAAATTTGAGGCTCACATTGCCACAGTGAAG 240
QY 1312 AGGAAGCTTTTACCGGAGGATATCTCTGAGAAAAGTTGAAAGATGGCAAAATGCTTGTG 1371
Db 241 AGGAAGCTTTTACCGGAGGATATCTCTGAGAAAAGTTGAAAGATGGTGAATGATGCTG 300
QY 1372 AGAGGAGCTGCTGCGAGGATTAACAACAAGATGATAATCGTTACAAATGCTGGGAGG 1431
Db 301 AGGAGGAGCTGCTGCGAGGATTAACAACAAGATGATAATCGTTACAAATGCTGGGAGG 360
QY 1432 GAGCACATTAACACATTAAGAGGCTGAAAAGGCTCGTACTTTGCTCAACAGATTCCTG 1491
Db 361 GGGCCCACTTGACATCAAAAGGCTGAGAAGCTCGTATTTGGTTAAACAGATCCAG 420
QY 1492 GAATGATAGATTTTGAACAACAAAATTTGCTGCAATGGAAAATGAAAGAGGAGG 1551
Db 421 GACTGCTAGATGTTTTAACCAACAAAATTTGAGCCTGGGAGGAGCAAGAGAGAA---AGG 477
QY 1552 ATTTTCACATATGCTGTTAGCCTTCTGCTCAATGCTTGATGAATATATGTTGCTGTC 1611
Db 478 AATTTCACATATGCTGTTAGCCTTCTGCTCAATGCTTGATGAATATATGTTGCTGTC 537
QY 1612 AGGAGAAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1671
Db 538 AGGAGAAAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1701
QY 1672 AAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1701
Db 598 AAGCTGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 627
```

RESULT 8

```
CA254411 1076 bp mRNA linear EST 25-SEP-2003
LOCUS SCBF4113G10.g FL4 Saccharum officinarum cDNA clone SCBF4113G10
5', mRNA sequence.
```

ACCESSION CA254411

VERSION CA254411.1

KEYWORDS GI:35342301

SOURCE Saccharum officinarum

ORGANISM Saccharum officinarum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum complex.

REFERENCE

1 (bases 1 to 1076)
Vettore, A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)

COMMENT

Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccen.br/facv.unesp.br
Plate: 113 row: G column: 10
Seq primer: T7 Promoter Primer.

FEATURES

source

1..1076

/organism="Saccharum officinarum"

/mol_type="mRNA"

/db_xref="taxon:4547"

RESULT 9

CA298359

LOCUS

DEFINITION

SCRFFL8035E05.g FL8 Saccharum officinarum cDNA clone

CA298359

667 bp

mRNA

linear

EST 26-SEP-2003

SCRFFL8035E05

/clone="SCBF4113G10"

/lab_host="DH10B"

/clone_lib="FL4"

/note="Organ: Developed inflorescence and rachis (20cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI; An unidirectional cDNA library generated from [Developed inflorescence and rachis (20cm-long)] cDNA was prepared from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at
http://suest.lad.ic.unicamp.br/public"

ORIGIN

Query Match 17.4%; Score 429.2; DB 6; Length 1076;
Best Local Similarity 79.3%; Pred. No. 3.8e-99;
Matches 548; Conservative 0; Mismatches 134; Indels 9; Gaps 3;

QY 1585 TGCTTGATGATATATGTTGTTGCTCAGGAGAAAGAGCAAGAGAGAGAGAGAGAGAGAGG 1644

Db 51 TCCGTGAAGAGTACATGATCGTTCCGAGGAGAAAGAGCTAGAGAGAGAGAGAGAGAGG 110

QY 1645 ATCAGAAAGAGCTCCAGGATCAGCTCAAGCGGAGAGAGAGCTTTGTACGATCAAAAC 1704

Db 111 ATCAGAAAGAGATCCAGATCACTCAAGCTGAGCAGAGAGAGAGCTGTATGGATCAAAAC 170

QY 1705 CCAGTCCATCCAGCCCTTAAGTACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1764

Db 171 CAAGTCCATCTAAGCCCAAG 230

QY 1765 ACCGAGGCTATCTCTGTTGGAGCCACCATGCAACCCCGAGAGAGAGAGAGAGAGAG 1824

Db 231 ACCGAGAGCTGTCTCTGTTGGAGCCCAATGCAAGCCCGAGAGAGAGAGAGAGAGAG 290

QY 1825 CAAAGTCTGTTGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1884

Db 291 CCAGAGAGCTGCTGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 347

QY 1885 GTAGAGGTTGGACATTTGCCGATGCTTATCAAGAGAGAGAGAGAGAGAGAGAGAGAG 1944

Db 348 GTAGAGGTTGGACATTTGCCGATGCTTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407

QY 1945 TACGTGAG 2004

Db 408 TGAAG 467

QY 2005 CGTCAG 2064

Db 468 CATCGGTCCTGCGGAG 524

QY 2065 AGACATTTACAGCAGTCAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2124

Db 525 AGACATTTGAGAGGCTCAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584

QY 2125 CAATGACTCTCTCTGCGCAAG 2184

Db 585 CCGTGACAG 644

QY 2185 AGCCAGAGAGT---AACATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2241

Db 645 AGCCAGAGGTTGCCAG 704

QY 2242 CCATCTATCTGCGGAG 2272

Db 705 CCGGTTACCTGCGGAG 735

RESULT 9

CA298359

LOCUS

DEFINITION

SCRFFL8035E05.g FL8 Saccharum officinarum cDNA clone

CA298359

667 bp

mRNA

linear

EST 26-SEP-2003

SCRFFL8035E05

```

5', mRNA sequence.
CA298359
CA298359.1 GI:36068563
EST.
Saccharum officinarum
Saccharum officinarum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
complex.
1 (bases 1 to 667)
Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
The libraries that made SUCSEST
Gener. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
http://www.bcccenter.fcav.unesp.br
Plate: 035 row: E column: 05
Seq primer: T7 Promoter Primer.
Location/Qualifiers
1. .667
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCRPL803505"
/lab_host="DH10B"
/clone_lib="FL8"
/notes="Organ: Developing inflorescence and rachis
(10cm-long); Vector: pSport1; Site 1: SalI; Site 2: NotI;
An unidirectional cDNA library generated from [Developing
inflorescence and rachis (10cm-long)]. cDNA was prepared
from polyA+ mRNA using SuperScript Plasmid System kit
(invitrogen). The double-strand cDNAs were fractionated
in a sepharose CL-2B 40cm-columns and fragments sizing
between 0.8 and 1.5 kb were directionally cloned into the
vector. Details of each source of RNA and library
construction can be obtained at
http://sucst.lad.ic.unicamp.br/public"

FEATURES
source
RESULT 10
BG487823/c
LOCUS
DEFINITION
PROPINGUUM CDNA, mRNA sequence.
ACCESSION
BG487823
VERSION
BG487823.1
KEYWORDS
EST.
SOURCE
Sorghum propinquum
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 617)
Cordonnier-Pratt,M.-M., Gingle,A., Sudman,M., Marsala,C. and
Pratt,L.H.
An EST database from Sorghum: floral-induced meristems
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmprratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: T7
High quality sequence start: 14
High quality sequence stop: 613
POLYA=NO.
Location/Qualifiers
1. .617
/organism="Sorghum propinquum"
/mol_type="mRNA"
/db_xref="taxon:132711"
/clone_lib="Floral-Induced Meristem 1 (FM1)"
/notes="Organ: Floral-Induced meristems; Vector:
pluscript II from Lambda Zap II; Site 1: XhoI; Site 2:
EcoRI; mature plants were placed in a growth chamber for
15 days with 16 hr darkness and 8 hr light (flowering is
induced by short-day conditions); 16 days after being
returned to the greenhouse under natural long days during
late April/early May, meristems were harvested the
library was made from poly-A RNA in the cloning vector

FEATURES
source
Query Match 16.7%; Score 411.8; DB 6; Length 667;
Best Local Similarity 77.2%; Pred. No. 1e-94;
Matches 515; Conservative 0; Mismatches 147; Indels 5; Gaps 1;

QY 56 CTCCTCCCTCTCTCGGCTCTCTGCAATCTGAGCTCCGATCGCGCGGACCCAGCCAG 115
DB 1 CACTCTCTCGCCCTCTGAGGCGGCTGATCGCACAAACTCTCCCGGCGACAGTCCGCGC 60

QY 116 AATCGCGCCCGCTCTCGCCCTCCCGCTCGAGAGAGCCGCGCGAGCGCGAGAGGC 175
DB 61 CCGTCCCGCCCGCGCGCGCGCGCGCTCGAGAGGCTGCGCGGCTGTGAAGAGGG 120

QY 176 -----CTAGTGTCTTCTCGCACTCTCGGATGAGTAGCGGTGAAGGACCAAGCTTCACCA 230
DB 121 TGTGCTGCGGTGTCTCCCTCTCGCTATGATGAGCGCTGTGAAGGACCAAGCTTCAGCAG 180

QY 231 ATGTGACGACATCGCATTCGTTCTACTGGAGCTCAATGTATTTGGGATGAGTTCGTT 290
DB 181 ATGTGACGACGCTGCGATTCCTCTCTGAGCTCAACGCTGATTTGGGATGAGTTCGTT 240

QY 291 GAGCCCCACACGACGAGGACAGATCTCTGAGCTTCGAGCTCCGAGGAGTCCCTGGAGGTC 350
DB 241 GAGCCCCACATGTTGAGGAGCCGAGATCTCTGAGCTTCGAGCTCGAGGAGTGCCTCGAGTC 300

QY 351 TACAGGCGGAGGTTCGACACGAGGCGAAACCGGAGCGCGCCAGCTGCCGGAAGGCCATGCC 410

```

```

Db 301 TACAGGAGGAAGGTTCGACGAGCCCAATCGGTGAGGGGCCAGTTTCGGGACGCCATTGCG 360
QY 411 GAGGCGGAGGAGAGCTCGCGGCAATCTGCTCAGCATGCGGAGAGCCGCCCTGACGCTT 470
Db 361 GAGGAGAGGCGGAACTCGCTGGCATCTGCTCGGCAATCGGCGAGCCACCGATACATGTT 420
QY 471 AGACAGTCAATCAGAACTTCATGGCTTAAGAGAGGAGTTGAATGCAATTCGTTCCGTAT 530
Db 421 AGACAGTCAATCAGAACTTCATGGTTTGAAGGAGGAATTTGAATGCGGATTGTCCCATAC 480
QY 531 TTGGAAGAAATCAGAAAGAAAAGGTGCAACGATGGAACCACTTTGTTTCATGTCATAGAG 590
Db 481 TTGGAAGAGATCAGAAAGAAAGAACTGAAGATGCGGCAATTTGTTGATGTTATAGAG 540
QY 591 CAGATTAAGAAAATTTTCGTTCTGAAATAAGGCCAGCGGATTTTGTTCCTTTAAAGTCCG 650
Db 541 CAAATTAAGAGGTTGCACTCTGAAATCAGGCTGAGGATTTTGCACCTTTGAGAAATTCCT 600
QY 651 GTTGATCAGTCTGACCTGTCATTAAGAAAGCTTTGATGAGTTGACGAGGACCTGGGATCC 710
Db 601 GTGGATCAGTCTGATCTGTCTTGAAGAAAGCTGAGGAGCTAATAAAGAGCTTACAATCC 660
QY 711 CTTTCA 717
Db 661 CTTTCA 667

```


Qy	797	T T T C A A G C A A A C A G T A T A T G A G G T G C A C C T A G C T T G G A C G A A G C T G A A G G A T C A A A G A A	856
Dd	61	C T T C A A A C A A A C A G T G C A T G A G T G C A C C T A G C T T G G A C G A A G C T G A A G G A T C A A A G A A	120
Qy	857	C C T G A C A A C A C T A C A A T T G A G A G C T T C T C T G C C G C A A A C A G A C T G C G T C A A A T G A A	916
Dd	121	C C T G A C A A C A C T T C A T C G A G A C T A G A C A T T G G C T G T G G A T A G A C T T C G T C A A A T A A A	180
Qy	917	G A T C C A A A G G A T G C A A A A G C T T C A A G A T T T G C T T C T A G C A T G C T C G A G C T A T G G A A T C T	976

977 CATGGATCTCCAATTGAAGAGCAGCAGATGTTTCAGAAATATCAATGCAATATTGTCG 103
QY
241 TATGNTACACCAATTGAAGAGCAGCAGATGTTTCAGAAATGTAACGTCGAATATTGTCG 300
Db
1037 TTCAGAAACAGAGATAAATCGAACCAACCCCTCTCCACAGATTTCTGAAATTAATGTCGA 1096
QY

301	CTCGAGCATGAAATAACGGACCCAAACCCCTGTGCGATCGACTTCTCAGCTATGTGGA	360
Db		
1097	ATCTGAGGTGTTAAGGCTTTGAACAACCTGAAGCAAGTAAAGATCAAAAGATCTTGTTTTAAA	1156
Qy		
361	AGCTGAGGTTTTAAGGCTTTGAGCAATTGAAGGAAGCAAAATGAAGACCTTGTTTTGAA	420
Db		
1157	AAAGAAAGCAGAACTAGAAGAGCATAGAAGAGCTGCTCATCTTGTGGCGAGAAAGTTA	1216
Qy		
421	AAAGAAATCAGAACTCGAGGAGCAGAAAGACCGCGCATCTAATTTGGCGAGNAGATA	480
Db		
1217	TGCAGAGGAGTTTAGCATTTGAAGCTTGAAGCTGGAGCTATTGATCCCTCAGTAGTACT	1276
Qy		
481	TTCAGATGAATTTAAACATTTGAGGCTATTGAGTCGGAGGCTATTGATCCTGCATTTGGTGCT	540
Db		

Qy	1277	TGAACAAA	1284	
Db	541	GGAACAAA	548	
RESULT 13				
CNS0A2G0				
LOCUS	CNS0A2G0	1986 bp	mRNA	linear
DEFINITION	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTPEH20ZC09 of Hormone Treated Callus of strain col-0 of			

ACCESSION	EX827777
VERSION	EX827777.1 GI:42459968
KEYWORDS	HTC; GSUT cDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eutrosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	Castelli, V., Aury, J.M., Jaillon, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1986)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage

COMMENT

The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.

- Web : www.genoscope.cns.fr
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)

TITLE Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2195)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr]

COMMENT The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full_length
<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES
Location/Qualifiers
1..2195
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTSL592C11"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..2195
/gene="At5g5230"

ORIGIN

Query Match 16.0%; Score 395.6; DB 3; Length 2195;
Best Local Similarity 55.7%; Pred. No. 2.2e-90;
Matches 846; Conservative 0; Mismatches 659; Indels 15; Gaps 4;
QY 222 CTTACACAGATGTCAGACATGCGATTGCTTCTACTGGAGCTCAATGATTTGGGAT 281
DB 146 CCTCATCTTTGGGAAATTAATCTTGTGCTACTTACTTGAAGAATTTGCAGGAATCTGGAT 205
QY 282 GAGTCCGTGAGCCACACAGCAGGACAGATGCTGCTGGAGCTCAGCAGGAGTGC 341
DB 206 GAAGTTGGTGAGAGTGATGATGAACGAGACAACTGCTTCTTCAGATAGCAGAGTGT 265
QY 342 CTGGAGGTCTACAGCGGAGGTGCGACCGCGGAAACCGGAGCGCGCCAGCTGCCGAAG 401
DB 266 CTTGACGTTTACAGAGAAAGTTCGACGAGCTGCGAAATCCGAGCTGAGCTTCTTCAA 325
QY 402 GCCATCCCGAGGCGAGGAGCTGCGCGGATCTGCTAGCCATGGCGAGCGCC 461
DB 326 ACCTTGTCAGATGCTAATGCTGAACCTTCAGGCTCACAATGCTCTCTGAGACAAAGC 385
QY 462 GTCCAGCTTAGACAGTCAATCAGAGCTTCATGGCTTAAGAGGAGTGAATGCAATT 521
DB 386 TT---AGTTGGCAATCCGGATAAGTCTTCAGGAACGATTAAAGAACAACTTGTGCAATA 442
QY 522 GTTCCGTATTGGAAGAAATGAAAGAAAGAGTTCGACGATGGAACAGATTGTTTCAT 581
DB 443 GCACCGGCTCTTGAAACAACTGTGGCAACAGAAAGAGAGAGTCCGAGAGTCTCTGAT 502
QY 582 GTCATAGCAGATTAAGAAATTTGCTCTGAAATAAGCCAGCCGAGTTTGTTCCTTT 641
DB 503 GTACATCAGATCAGAGATATGAGAGATAT---TGCTGGAGTTTGCAATGAG 559
QY 642 AAGTTCGGTTCATGCTGACCTGTCTATTAAGAAAGCTTCATGATGACGAGGAC 701
DB 560 GTTCCCTATAGTCGATGATGATTTGTCACTGAAGAAATTAGACGATTTCCAGGCCAA 619
QY 702 CTGGAATCCCTTCAGAGGAGAGAGCGATCGGCTAAAGCAAGTATAGAACATTTGAAT 761

RESULT 15

DB 620 CTCCAAGAGCTCCAGAGAGAGAGAGTGCAGAGGCTGCGCAAGGTTTAGAGTTTGTGAGT 679
QY 762 TCTTTGCAATCTTATGTGAGGTGCTTGGCATAGATTTCAAGCAAAAGTATATATGAGGTG 821
DB 680 ACTGTTCAATGATCATGTGCTGCTTCTTGGTTTGGATTTCTTAAGCACCCTCAGCAAGTT 739
QY 822 CACCTAGCTTGGAGCAAG-----CTGAAGGATCAAAAGAACTTGAAGCAACATCAATT 875
DB 740 CATCGAGCTTAGATGAAGATACCAAGTGTCCAGTCTAAGAGCATTAGCAATGAGACTCTT 799
QY 876 GAGAGCTTGTGCTGCGCAAAAGATCGCTGTAATGAAGATCAAAAGGATGCAAAAG 935
DB 800 TCAAGTTGGCTAAGACCGCTCTTGACTCTTAAGAGATGATGAAGACACAGACTTCAAAAG 859
QY 936 CTTCAAGATTTTGTCTTAGCATGCTCGAGCTATGGAATCTCATGGATCTCCACTTGA 995
DB 860 CTTCAAGAGCTGCTACTCAGCTAATGACCTGTGGAATCTGATGATCTCTCTGATGAG 919
QY 996 GAGCAGCAGATGTTTTCAGAAATATAACATGCAATATTGCTGCTTCAGAAACAGAGATACT 1055
DB 920 GAAAGAGAGCTTTTGTATCATGTTACCTGTGACATTTTCATCTTCACTGATGAGTCACT 979
QY 1056 GAACCAACACCTCTCCAGATTTCTGAAATTTGTCGAATCTGAGGTGTTAAGGCTT 1115
DB 980 GTGCCAGGTGCTCTTGGCCGCTGATTTGATGAGCAGGCTGAGGTGGAAGTTGATAGGCTT 1039
QY 1116 GAACCACTGAAGCAAGTAAGATGAAGATCTTGTGTTTAAAGAAAGAGCAACATAGAA 1175
DB 1040 GACCACTGAAGCTAGCGAATGAAGAAATTTGCTTCAAGAGCAATCTGAGCTTGA 1099
QY 1176 GAGCATAGAAGACGTGCTCATCTTGTGGCGAGGAAGGTATTGACAGAGGATTTAGCAT 1235
DB 1100 GAGATATATGCTCGTCCCATGTAGAAGTTAACCCGGAATCTGCTGAGAGAAATCATG 1159
QY 1236 GAAGCTATTGAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTCAC 1295
DB 1160 TCGCTGATTGATTCTGGAACGTTGAGCTTACTGAAATTTTGGCAGACATGATAGCCAG 1219
QY 1296 ATTGCAACAGTCAAGAGAGAGGCTTTTACCGGAAGGATATTCTTGAAGAGTTGAAGA 1355
DB 1220 ATATCAAGGCTTAAGAAAGAGCATTTAGTAGAAAGATATATTGGACCGAGTCGAGAA 1279
QY 1356 TGGCAAAATGCTTTGTAAGAGAGAACCTGGCTGGAAGATTACAAACAAAGATCAATCTG 1415
DB 1280 TGGATGTGAGCTTTGTGAGAGAGAGAGCTGGCTAGAAGACTACAATCGGGATCAGAACAG 1339
QY 1416 TACAATGCTGGAGGGAGCAGCATCTAACATAAGAGGGCTGAAGAGGCTGCTACTTTG 1475
DB 1340 TACAGCGCAAGCAGAGGTGCACATTTGAATCTCAAGAGAGCTGAGAAAGCTCGGATTCG 1399
QY 1476 GTCAACAGATTCTGGAATGCTAGATGTTTGAAGAAACAAATTTGCTCATGGAAAAAT 1535
DB 1400 GTTAGCAGATTCTCGCCATGTTGACACATTTAGTTGCCAAGACCGGGCTTGGAGAA 1459
QY 1536 GAAACAGAAAGGAGGATTTCAATATGATGTTGTTAGCTTTTCTGCTTCAATGCTGATGAA 1595
DB 1460 GAACACAGCA---TGTCCTTTGCTTACGATGCTGTTCTCTCTGCTAGCTATGCTAGACGAG 1516
QY 1596 TATATGTTGCTGCTCAGAGAAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1655
DB 1517 TACGGTATGCTTAGGCAAG 1576
QY 1656 CTCAGGATCAGCTCAAGCGGAGGAGAGAGGCTTTGTACGGATCAAAACCCAGTCCATCC 1715
DB 1577 GTTCAAGAACACCCACAGTGAAGCAAGAAATCTGCTTTAGCAACGAGCCCTGCA 1636
QY 1716 AAGCCCTTAAGTACAAAGAA 1735
DB 1637 AGACCGGTCACTGCTAAGAA 1656

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	100.4	4.1	267	4	US-09-313-294A-843	Sequence 843, Appl Sequence 950, App
2	92	3.7	298	4	US-09-313-294A-950	Sequence 14, Appl
3	73.8	3.0	7218	1	US-08-232-463-14	Sequence 6877, App
4	59	2.4	296	4	US-09-313-294A-6877	Sequence 2473, App
5	47.6	1.9	3763	4	US-09-919-039-243	Sequence 640, App
6	47.6	1.9	5185	4	US-09-976-594-640	Sequence 2, Appli
7	46	1.9	1425	2	US-08-578-551-2	Sequence 2, Appli
8	46	1.9	1425	2	US-09-190-982-2	Sequence 3, Appli
9	46	1.9	1425	3	US-09-408-257-2	Sequence 3, Appli
10	43.4	1.8	16550	4	US-08-916-421B-3	Sequence 3, Appli
11	43.4	1.8	16550	4	US-09-692-570-3	Patent No. 5231168
12	43.2	1.8	3095	6	5231168-1	Sequence 14, Appl
13	43	1.7	7218	1	US-08-232-463-14	Sequence 1, Appl
14	42.2	1.7	1230025	4	US-09-198-452A-1	Sequence 67, Appl
15	42	1.7	2082	4	US-09-818-780-67	Sequence 2, Appli
16	42	1.7	2277	1	US-08-676-967-2	Sequence 2, Appli
17	42	1.7	2277	1	US-08-676-974-2	Sequence 2, Appli
18	42	1.7	2277	2	US-09-098-487-2	Sequence 69, Appl
19	41.8	1.7	798	4	US-09-724-797-69	Sequence 6992, Ap
20	41.6	1.7	786	4	US-09-252-991A-6992	Sequence 3643, Ap
21	41.6	1.7	828	4	US-09-252-991A-3643	Sequence 3817, Ap
22	41.6	1.7	1458	4	US-09-252-991A-3817	Sequence 3864, Ap
23	41.6	1.7	1779	4	US-09-252-991A-3864	Sequence 7043, Ap
24	41.6	1.7	2412	4	US-09-252-991A-7043	Sequence 7084, Ap
25	41.6	1.7	2472	4	US-09-252-991A-10200	Sequence 10200, A
26	41	1.7	855	4	US-09-252-991A-10200	Sequence 14937, A
27	41	1.7	978	4	US-09-252-991A-14837	

; Sequence 950, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 950
; LENGTH: 298
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700550124H1
US-09-313-294A-950

Query Match 3.7%; Score 92; DB 4; Length 298;
Best Local Similarity 58.3%; Pred. No. 1.3e-15;
Matches 161; Conservative 0; Mismatches 115; Indels 0; Gaps 0;
QY 902 ACTGCGTGAATGAAGATGCAAGGATGCAAAAGCTTCAAGATTTTCTTAGCATGCT 961
Db 17 ACTTGAAGAGAAAATGAAGAGTTTGCAGAGTTTGCAGAGTTTCAAGCTTAGCTTCCAGCTATC 76
QY 962 CGAGCTATGGATCTCATGGATCTCCACTTGAAGAGCAGCAGATGTTTTCAGAAATATAAC 1021
Db 77 TGATCTTTGGAATTAATGGATCTCTCTGTGAGGAAACAGCAACCTTTTCGTACAGTAAC 136
QY 1022 ATGCAATATGCTGCTTCAAGACAGAGATACTGAACCAACACCTCTCCACAGATTT 1081
Db 137 ATGCATAATGCTCCTCAACATTGGATGATGTAGTATTCAGGAGCTCTGGCTCTTGATGT 196
QY 1082 CCGTAATATGTCGAATCTGAGTCTTAAGGCTTCAACAACTGAAAGCAAGTAAGATGAA 1141
Db 197 AATTCAGACGGCGAAGCTCGAAGTTGAAGGCTTGAATGAAGCTAGTAAAGCTAGTAGTAA 256
QY 1142 AGATCTTGTGTTTAAAAAGAAAGCAGAACTAGAAGA 1177
Db 257 GGATATGTCATTCAAGAAACAGACTGAACTTGAAGA 292

RESULT 3
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 INMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14
Query Match 3.0%; Score 73.8; DB 1; Length 7218;
Best Local Similarity 4.7%; Pred. No. 1.2e-09;
Matches 18; Conservative 231; Mismatches 138; Indels 0; Gaps 0;
QY 1269 CTAGTACTTGACAAATTTGAGCTCACATTCGAACAGTGAAGAGAGCTTTTAGCCGG 1328
Db 1466 CAAGTAGTTAAGAGATAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRR 1407
QY 1329 AAGGATATCTTGTGAGAAATTTGAAAGATGGCAAAATGCTTGTGAAGAGAAAGCTGGCTG 1388
Db 1406 RRR 1347
QY 1389 GAAGATTACAAAGATGATACTGTTACAACTCTGGAGGGAGGACACTTCAACTA 1448
Db 1346 RRR 1287
QY 1449 AAGAGGGCTGAAAGGCTCGTACTTTGGTCAACAAGATCTCTGGAATGGTAGATGTTG 1508
Db 1286 RRR 1227
QY 1509 AGAACAATAATTTGCTGATGCAAAAAATCAACGAGAAAGAGGATTTTCAATATGATG 1568
Db 1226 RRR 1167
QY 1569 GTTAGCCCTTTCGTAATGCTTGTGAATATATGTTCTGTCAGGAGAAAGACAAG 1628
Db 1166 RRR 1107
QY 1629 AAGAAGAGACAAGGGATCAGAAG 1655
Db 1106 RRR 1080

RESULT 4
US-09-313-294A-6877
; Sequence 6877, Application US/09313294A
; Patent No. 6476212
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.
; APPLICANT: Ito, Laura Y.
; APPLICANT: Sherman, Bradley K.
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
; FILE REFERENCE: PL-0017 US
; CURRENT APPLICATION NUMBER: US/09/313,294A
; CURRENT FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 7600
; SOFTWARE: PERL Program
; SEQ ID NO 6877
; LENGTH: 296
; TYPE: DNA
; ORGANISM: Zea mays

[illegible]

```

RESULT 5
US-09-919-039-243
; Sequence 243, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 243
; LENGTH: 3763
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 346209.3
US-09-919-039-243

```

	Query Match	1.9%	Score 47.6;	DB 4;	Length 3763;
	Best Local Similarity	50.4%;	Prod. No. 0.014;		
	Matches 116;	Conservative	0;	Mismatches 114;	Indels 0; Gaps 0;
Qy	313	GGATGCTGCTGGAGTCTCAGCAGAGATGCCCTGGAGGCTCTACAGGCGGAAGTTCGACACAGG	372		
Db	316	GGGCCCCAGCTGGGCCGGAAGAGAGAGAGCTTCAGGCTGCCCTGSCCAGGCGAGAAGACG	375		
Qy	373	CGAACCAGGAGCGCGCCCGAGCTGCGGAAGGCCATCGCCGAGGGCGGAGCAGAGCTCGCCG	432		
Db	376	AGGGTGGGGCCCGGCGCCAGCTGTCTGAAATCCCTGCGGGAGGCTCAAGCAGCCCTGGCCG	435		
Qy	433	GCATCTGCTCAGCCATGGCGAGCGCCCGCTGCAGTTAGACAGTCAAAATCAGAAGCTTC	492		
Db	436	AGGCCCAGAGGACCTGAGTCTGNGCGTGTGGCCAGGACCAAGCGCGAAGACGACGCC	495		
Qy	493	ATGCTTTAAGAGAGAGGTTGAATGCAATGTTTCCGTATTTTGGAGAAGATG	542		
Db	496	GGGACCTGGGCGAGAGAGCTGAGGGCGCTGCGGGGCGAGCTGGAGACACAG	545		

RESULT 6
US-09-976-594-640
: Sequence 640, Application US/09976594

```

; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Suchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH S
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 640
; LENGTH: 5185
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 346209.3
; US-09-976-594-640

```

	Query Match	1.9%;	Score 47.6;	DB 4;	Length 5185;
	Best Local Similarity	50.4%;	Pred. No. 0.018;		
	Matches 116;	Conservative	0;	Mismatches 114;	Indels 0; Gaps 0;
Qy	313	GGATGCTGCTGGAGCTCGACGACGAGTGCCTGGAGTCTACAGGCGGAAAGTTCGACCAGG	372		
Db	1738	GGGCCCAGCTGGGCCGGAAGGAGGAGAGCTGCAGGCTGCCCTGGCCAGGCGCAAGAAGCG	1797		
Qy	373	CGNAACCGGAGCCGCGCCCAAGCTTCGCGAAGGCCATCGCCGAGGCGGAGGAGGAGCTCGCCG	432		
Db	1798	AGGGTGGGGCCCGGGCCCAAGCTGCTGAAATCCCTCGGGAGGCGTCAAGCAGGCCCTGGCCG	1857		
Qy	433	GCATCTGCTCAGCCATGCGGCGAGCCGCCGTCACGTTAGACAGTCAAAATCAGAAAGCTTC	492		
Db	1858	AGGCCACAGAGGACCTGGAGTCTGAGCGTGTGCCAGGACCAAGCGCGGAGMACGACGCGCC	1917		
Qy	493	ATGGCTTAAGAGAGGAGTGGAAATGCAATGTGTTCCGTATTTGGAGAAATG	542		
Db	1918	GGGACCTGGGCGGAGGAGCTGGAGGCGCTCGGGGGCGAGCTGGAGGACAG	1967		

RESULT 7
US-08-578-551-2/c
; Sequence 2, Application US/08578551
; Patent No. 5854050
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Christgau, Stephan
; APPLICANT: Andersen, Lene N.
; APPLICANT: Kofod, Lene V.
; APPLICANT: Kauppinen, Sakari M.
; APPLICANT: Nielsen, Jack B.
; APPLICANT: Dambmann, Claus
; TITLE OF INVENTION: An Enzyme with Protease Activity
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5854050o No. 5854050disk of No. 5854050th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,551
; FILING DATE: 01-FEB-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

RESULT 9
US-09-408-257-2/c
; Sequence 2, Application US/09408257
; Patent No. 6150905
; GENERAL INFORMATION:
; APPLICANT: Dalboge, Henrik
; APPLICANT: Christgau, Stephan
; APPLICANT: Andersen, Lene N.
; APPLICANT: Kofod, Lene V.
; APPLICANT: Kauppinen, Sakari M.
; APPLICANT: Nielsen, Jack B.
; APPLICANT: Dammann, Claus
; TITLE OF INVENTION: An Enzyme w
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:

```


ADDRESSSEE: No. 61909050 No. 61909050disk of No. 61909050th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/408,257
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/578,551
FILING DATE: 01-FEB-1996
APPLICATION NUMBER: DK 0811/93
FILING DATE: 06-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 95/02044
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4006.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1425 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Aspergillus aculeatus
US-09-408-257-2

Query Match	1.9%	Score 46;	DB 3;	Length 1425;
Best Local Similarity	47.3%;	Pred. No. 0.022;		
Matches 139;	Conservative 0;	Mismatches 155;	Indels 0;	Gaps 0;
Qy	133	CGCCCTCCCGCTCGACGAGACCGCGCGGAGCGCGAAGAGGCTAGTGTCTTCGCA	192	
Db	578	CGCCCTCGCTGCTGGCCGCTCACCGTCCGACGGTGACCTTGTCTTGAGACGTGCGCGC	519	
Qy	193	TCGCGATAGTAGCGGCTCAAGACACAGCTTACACAGATGTCCACGACATGCGATTCG	252	
Db	518	TGGCGAGCTGCCGCTCCCGTAGAGATGCTCAGCTGTAGCCGACAGCTTCTTGGCGG	459	
Qy	253	TTCTACTGGAGCTCAATGTGATTGGGATGAGGTCCGTGAGCCCGACACGACGAGGACA	312	
Db	458	TGTCGCGGGGTGTAGATGCTGTGGCCGCTGCCGAGGAGCGCCAGCTCCGAGGAGA	399	
Qy	313	GGATGCTGCTGGAGCTCCAGCAGAGATGCCCTGGAGGTCTACAGCGGAAAGGTTCGACCAGG	372	
Db	398	AGACCCAGAGATCGCGGATCCGGTGTGMAATGCCAGGTTCAAGGTCCAGCTGCCACCG	339	
Qy	373	CGAACCGAGCGCGCCCGCTGCGGAAGGCCATCGCCGAGGGCGAGGCAGAC	426	
Db	338	TGACGGGGGTTCAGCTAGGAGACGTGCTTTGGCTCGCGGGTGGTCAACGACGTGC	285	

RESULT 10
US-08-916-421B-3/c
; Sequence 3, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:

```

/ APPLICANT: Bult et al.
/ TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii
/ Patent No. 6503729
/ TITLE OF INVENTION: jannaschii
/ FILE REFERENCE: PB275
/ CURRENT APPLICATION NUMBER: US/08/916,421B
/ CURRENT FILING DATE: 1997-08-22
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 3
/ LENGTH: 16550
/ TYPE: DNA
/ ORGANISM: Methanococcus jannaschii
/ US-08-916-421B-3

```

Query Match	1.8%	Score	43.4	DB	4	Length	16550
Best Local Similarity	47.9%	Pred. No.	0.55				
Matches	125	Conservative	0	Mismatches	136	Indels	0
Gaps	0						
Qy	1120	AACTGAAACGACGTAGATGAAGATCTTGTTTAAAAAAGAAAGCAGAACTAGAAGAGC	1179				
Db	4285	ATCTTAAATCTATTTTGTAGTATGTTCTTGAAATAAAGAAAGATAAAGATTTAGAAATG	4226				
Qy	1180	ATAGAGAGCGTCTCATCTTGTTGGCGAGGAGTTATGCAGAGGAGTTTAGCATTTGAAG	1239				
Db	4225	AATTAAGACGGTATTTTCTGAGTTTATTTGTTTTTCAAGGAGTATATTGAAGAAT	4166				
Qy	1240	CTATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTCACATTG	1399				
Db	4165	ATATTGAGTTGCTACAAAGATATTAAACATAAAGAAAGAAATTTACTAATGAATTC	4106				
Qy	1300	CACACGTGAAGAGGAGCTTTTAGCCGSAAGATATTTCTTGAGAAAGTTTGAAGATGCG	1359				
Db	4105	ATAAAGAGTTGCTGTAACCTGTTTTTGGATTTGAAGATGTTAAGGATGTTTGAAGAAAAAG	4046				
Qy	1360	AAATGCTTGTGAAGAGGAAG	1380				
Db	4045	ATAAAGTAGTTTCAGAAAAAG	4025				

RESULT 11

US-09-692-570-3/c
; Sequence 3, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, M
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 16550
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
US-09-692-570-3

	Query Match	1.8%	Score 43.4;	DB 4;	Length 16550;
	Best Local Similarity	47.9%;	Pred. No. 0.55;		
	Matches 125;	Conservative	0;	Mismatches 136;	Indels 0; Gaps 0;
QY	1120	AACTGGAAGCAAGTAAGATCAAAAGATCTCTTTTAAAAAAGAAAGCAGAACTAGAAAGC	1179		
DB	4295	ATCTTAAATCTATTTTGGATGTGTTCTTGAAATAATAAAAGATAAGATTTCGAAATG	4236		

QY 1180 ATAGAAGCGTCTCATCTTGTGGCGAGGAGTTATGCAGAGGATTTAGCATTTGAG 1239
|||
Db 4225 AATAAACAAGGATTTCTGAGTTTATTATTTTCAAGAGGATATATTGAAGAAT 4166
|||
QY 1240 CTATTGAAGTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTCAATTTG 1299
|||
Db 4165 ATATTGAAGTTGCTACAAAGATATTAAACATAAAGAAATTTACTAATGATTTCT 4106
|||
QY 1300 CAACAGTGAAGAGGAGCTTTTAGCCGGAAGATATTTCTTGAGAAAGTTGAAAGATGCC 1359
|||
Db 4105 ATAAAGAGTTCGTAAACATTTGTTTGGATTTGAAGATGTTAAGAGTGTAAAGAAAAAG 4046
|||
QY 1360 AAAATGCTTGTGAAGAGGAG 1380
|||
Db 4045 ATAAAGTAGTTCAAGAAAAAG 4025
|||
RESULT 12
5231168-1
; Patent No. 5231168
; APPLICANT: DZIEGIEL, MORTEN, BORRE, MARTIN, JEPSEN, SOREN;
; VUUST, JENS, RIENECK, KLAUS, WIND, ANNETTE, JAKOBSEN, PALLE H.
; TITLE OF INVENTION: MALARIA ANTIGEN
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; FILING DATE: 18-SEP-1989
; SEQ ID NO: 1:
; LENGTH: 3095
5231168-1
Query Match 1.8%; Score 43.2; DB 6; Length 3095;
Best Local Similarity 43.9%; Pred. No. 0.21;
Matches 186; Conservative 0; Mismatches 238; Indels 0; Gaps 0;
QY 1142 AGATCTTGTGTTTAAAAAGAAAGCAGAACTAGAAGAGCATAGAAGCGTGCTCATCTTGT 1201
|||
Db 1047 AGAATTTGTTGAAATTTGAGGAGTTTCTTGAACCAATCAAAATTAACGAAATTTCAAGA 1106
|||
QY 1202 TGGCGAGGAGGTATGCGAGAGGATTTAGCATTTGAAGCTATTGAAGCTGGAGCTATTGA 1261
|||
Db 1107 AATTATGAAGATGATTAAGGTGCAATATTCAGCATGAAATAGTAGAAGTAGAAGAAAT 1166
|||
QY 1262 TCCCTCAGTACTTGAACAAATTTGAAGCTCACATTTCAACAGTGAAAGAGGAGCTTT 1321
|||
Db 1167 ACTTCAGAGATGATTAATAATGAAAGTTGAAATGAAATAGTAGAAGTTGAAGAAAT 1226
|||
QY 1322 TAGCCGGAAGGATATCTTTGAGAAAGTTGAAAGATGGCAAAATGCTTGTGAAGAGGAAAGC 1381
|||
Db 1227 TCTACCAAGATTAATAATGAAAGGTTCAACATGAAATAGTAGAGGTTGAAGAAATTTCT 1286
|||
QY 1382 CTGGCTGGAAGATTACAAAGATGATAATCGTTTACAATGCTGGAGGGGAGGACATCT 1441
|||
Db 1287 ACCAGAAGATGATAATAATGAAAGATTTGAACATGAAATAGTAGAAGTTGAAGAAATTTCT 1346
|||
QY 1442 AACACTAAAGAGGCTGAAAGGCTCGTACTTTTGTCTCAACAGATTTCTGGAATGGTAGA 1501
|||
Db 1347 ACCAGAGATAAATAATGAAAGGTTCAACATGAAATAGTAGAGTTGAAGAAATTTCTACC 1406
|||
QY 1502 TGTTTTGAGAACAAATAATTTCTGCAATGGAATAATGAACGAGAAAGGAGGATTTTACATA 1561
|||
Db 1407 AGAAGATAAATAATGAAAGGTTGAACATGAAATAGTAGAAGTTGAAGAAATTTTACCAGA 1466
|||
QY 1562 TGAT 1565
|||
Db 1467 AGAT 1470
|||

RESULT 13
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367

; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14
Query Match 1.7%; Score 43; DB 1; Length 7218;
Best Local Similarity 2.9%; Pred. No. 0.42;
Matches 4; Conservative 99; Mismatches 34; Indels 0; Gaps 0;
QY 10 TTTCAAAATCACACTACACTCTCCGTCGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 69
|||
Db 1060 TTGCGATVY 1119
|||
QY 70 CGCCTCTCTCGCATCTGAGGCTCGGATCGCGGCGAGCCCGAGCAGCAATCGCGCGCCCG 129
|||
Db 1120 YV 1179
|||
QY 130 TCTCGCCCTCCCGCTC 146
|||
Db 1180 YV 1196
|||
RESULT 14
US-09-198-452A-1
; Sequence 1, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A

[illegible]

```
; LOCATION: (690001)..(705000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (705001)..(720000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (720001)..(735000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (735001)..(750000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (750001)..(765000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (765001)..(780000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (780001)..(795000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (795001)..(810000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (810001)..(825000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (825001)..(840000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (840001)..(855000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (855001)..(870000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (870001)..(885000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (885001)..(900000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
; LOCATION: (900001)..(915000)
; OTHER INFORMATION: n=a or c or g or t
; NAME/KEY: misc feature
```

```
Query Match      1.7%; Score 42.2; DB 4; Length 1230025;
Best Local Similarity 54.1%; Pred. No. 18;
Matches 86; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 1142 AGATCTTGTGTTTAAAAAAGACAGACTAGACAGCATAGACAGCTGCTCATCTTGT 1201
Db 348227 AGAAATGCTTCAATAACATCCCTTACCCAAATTAAGACCAGAACTGGGGCTCTTTT 348286

QY 1202 TGGCGAGGAAGTTATGACAGAGAGTTAGCATTTGAAGCTATTGAAGCTGAGCTATTGA 1261
Db 348287 TTCTCAATTAGATTTTGTATGAGACTTCGAGAACTTAGCTAAAGAAATACGACTCTGTTGA 348346

QY 1262 TCCCTCAGTACTTGAACAAATTTGAAGCTCACATTCG 1300
Db 348347 GCCTAAAGTAGAATTTCTGAAGGGAATAACATAGC 348385
```

```
RESULT 15
US-09-818-780-67
; Sequence 67, Application US/09818780
; Patent No. 6677146
; GENERAL INFORMATION:
; APPLICANT: McHenry, Charles
; TITLE OF INVENTION: NOVEL THERMOPHILIC POLYMERASE III HOLOENZYME
; FILE REFERENCE: 1794.0030004
; CURRENT APPLICATION NUMBER: US/09/818,780
; CURRENT FILING DATE: 2001-03-28
```

```
; PRIOR APPLICATION NUMBER: US 60/192,736
; PRIOR FILING DATE: 2000-03-28
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 2082
; TYPE: DNA
; ORGANISM: Thermus thermophilus
US-09-818-780-67
```

```
Query Match      1.7%; Score 42; DB 4; Length 2082;
Best Local Similarity 48.0%; Pred. No. 0.36;
Matches 120; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 260 GGAGCTCAATGTGATTTGGGATGAGGTGCGGTGAGCCCGACACGACGAGGGACAGGATGCT 319
Db 1488 GGACCGCTTGAGAACGTTGGAGAGCTCTCAGGGCGGCCCAAGGAGGCGGAGGACCTGCA 1547

QY 320 GCTGGAGCTCGAGCAGGAGTGCCTGGAGGTCTACAGGCGGAAGGTTCGACCAGGCGAACCG 379
Db 1548 GGACTTCTTGACCGGGTGGCCCTCACCGCCAAGCGGAGGAGGCCCGCGAGGGAGGG 1607

QY 380 GAGCGGCGCCAGCTGCGGGAAGGCCATCGCGAGGGCGAGGAGCTGCGCCGCACTG 439
Db 1608 GAGGTGCGCCCTCATGACCCCTGCACAACGCGCAAGGGGCTGGAGTTCCCGTGTCTTCCT 1667

QY 440 CTCAGCCATGGCGAGCGGCCCGTGCAGCTTAGACAGTCAATCAGAAGCTTCATGCGCTT 499
Db 1668 CGTGGGGTGGAGGAGGGGCTTCTGCCCCACCCGCAACTCGGTGAGCACCCCTCGAGGGCCT 1727

QY 500 AAGAGAGGAG 509
Db 1728 GGAGGAGGAG 1737
```

```
Search completed: November 20, 2004, 00:17:31
Job time : 212 secs
```

Result No.	Score	Query		DB	ID	Description
		Match	Length			
c	1	2468	100.0	2468	16	US-10-619-685-1 Sequence 1, Appli
	2	1467.6	59.5	2639	16	US-10-425-114-21434 Sequence 21434, A
	3	1439	58.3	2703	18	US-10-425-115-68593 Sequence 68593, A
	4	1021.2	41.4	4574	16	US-10-619-685-3 Sequence 3, Appli
	5	674.4	27.3	784	17	US-10-437-963-55930 Sequence 55930, A
	6	611.6	24.8	709	17	US-10-437-963-81367 Sequence 81367, A
	7	563	22.8	2045	17	US-10-437-963-98451 Sequence 98451, A
	8	506.6	20.5	530	17	US-10-437-963-32360 Sequence 32360, A
	9	467.8	19.0	2150	16	US-10-424-599-126041 Sequence 126041, A
	10	425.6	17.2	1680	17	US-10-437-963-77259 Sequence 77259, A
c	11	389.8	15.8	2811	17	US-10-437-963-88884 Sequence 88884, A
	12	389.6	15.8	2712	16	US-10-424-599-111490 Sequence 111490, A

Db CCTCTCTCGGCTCTCTCGCATCTGAGCTCCGATCGCGGCGACCCAGCCGAATCC 120
QY GCGGCCCGCTCTGCGCCTCCCGCTCGAGAGAGACCGCGCGAGCGGCGAAGGCGCTAGT 180
Db GCGGCCCGCTCTGCGCCTCCCGCTCGAGAGAGACCGCGCGAGCGGCGAAGGCGCTAGT 180
QY GTTCTTCGACCTCGCATGAGTACGCGGCTGAAGGACAGCTTCACACAGATGTCGACGA 240
Db GTTCTTCGACCTCGCATGAGTACGCGGCTGAAGGACAGCTTCACACAGATGTCGACGA 240
QY CATGCGATTTCGCTTACTGAGGCTCAATGTGATTTGGGATGAGGTCCGTCGAGCCCGACA 300
Db CATGCGATTTCGCTTACTGAGGCTCAATGTGATTTGGGATGAGGTCCGTCGAGCCCGACA 300
QY CGACGAGGACAGGATGCTGTCGAGCTCGAGCAGAGTGCTTGAGGTCTACAGGCGGA 360
Db CGACGAGGACAGGATGCTGTCGAGCTCGAGCAGAGTGCTTGAGGTCTACAGGCGGA 360
QY AGGTGACCGAGCGAAGCGGAGCGCGCCAGCTCGGAGCGCCATCCGAGGCGGAGG 420
Db AGGTGACCGAGCGAAGCGGAGCGCGCCAGCTCGGAGCGCCATCCGAGGCGGAGG 420
QY CAGAGCTCGCGGCACTGCTCAGCCATGCGGCGAGCGCGCCGTCGACGTTAGACAGTCAA 480
Db CAGAGCTCGCGGCACTGCTCAGCCATGCGGCGAGCGCGCCGTCGACGTTAGACAGTCAA 480
QY ATCAGAGCTTCATGCTTAAGAGAGGAGTGAATGCAATTTGTCGATTTTGGAGAAA 540
Db ATCAGAGCTTCATGCTTAAGAGAGGAGTGAATGCAATTTGTCGATTTTGGAGAAA 540
QY TGAAGAGAGAGAGTTCGACGATGGAACCGAGTGTTCGATGCTATAGCAGATTAAGA 600
Db TGAAGAGAGAGAGTTCGACGATGGAACCGAGTGTTCGATGCTATAGCAGATTAAGA 600
QY AAATTTGCTGTAAGAGCGCAGCGGATTTGTTCCCTTTAAAGTTCCGGTTGATCAGT 660
Db AAATTTGCTGTAAGAGCGCAGCGGATTTGTTCCCTTTAAAGTTCCGGTTGATCAGT 660
QY CTGACCTGTCAATTAAGAGAGTTCGATGATGAGTGAACGAGACCTGGAATCCCTTCAGAAAG 720
Db CTGACCTGTCAATTAAGAGAGTTCGATGATGAGTGAACGAGACCTGGAATCCCTTCAGAAAG 720
QY AGAAGAGAGTTCGATGATGAGTGAACGAGACCTGGAATCCCTTCAGAAAG 780
Db AGAAGAGAGTTCGATGATGAGTGAACGAGACCTGGAATCCCTTCAGAAAG 780
QY AGGTGCTTGGCATAGATTTCAAGCAAAACAGTATATGAGGTGCACCTAGCTTGGACGAAG 840
Db AGGTGCTTGGCATAGATTTCAAGCAAAACAGTATATGAGGTGCACCTAGCTTGGACGAAG 840
QY CTGAAGGATCAAGAACTGAGCAACA CTACATTAAGAGGCTTGCCTGTCGCGCAACA 900
Db CTGAAGGATCAAGAACTGAGCAACA CTACATTAAGAGGCTTGCCTGTCGCGCAACA 900
QY GACTGCTGTAATGAGATCCAAAGGATGCAAAAGCTTCAGATTTTCTGCTTACGATGC 960
Db GACTGCTGTAATGAGATCCAAAGGATGCAAAAGCTTCAGATTTTCTGCTTACGATGC 960
QY TCGAGCTATGGAATCTCATGATATCTCCATTTGAAGAGCAGAGATGTTTCAAGATATAA 1020
Db TCGAGCTATGGAATCTCATGATATCTCCATTTGAAGAGCAGAGATGTTTCAAGATATAA 1020
QY CATGCAATTTGCTGCTTCAGAACAGAGATTAAGTGAACCAACCTCTCCACAGATT 1080
Db CATGCAATTTGCTGCTTCAGAACAGAGATTAAGTGAACCAACCTCTCCACAGATT 1080
QY TCCTGAAATTTGCTGCTTCAGAACAGAGATTAAGTGAACCAACCTCTCCACAGATT 1140
Db TCCTGAAATTTGCTGCTTCAGAACAGAGATTAAGTGAACCAACCTCTCCACAGATT 1140
QY AAGATCTTGTGTTTAAAGAGAGAGATTAAGTGAACCAACCTCTCCACAGATT 1200

Db AAGATCTTGTGTTTAAAGAGAGAGATTAAGTGAACCAACCTCTCCACAGATT 1200
QY TTGCGGAGGAGGTTATGCGAGAGGTTTACATTTGAGAGCTATTGAACTCGAGCTATTG 1260
Db TTGCGGAGGAGGTTATGCGAGAGGTTTACATTTGAGAGCTATTGAACTCGAGCTATTG 1260
QY ATCCCTCACTAGTACTTTGAAACAAATTTGAAGCTCACATTTGCAACAGTGAAGAGGAGCTT 1320
Db ATCCCTCACTAGTACTTTGAAACAAATTTGAAGCTCACATTTGCAACAGTGAAGAGGAGCTT 1320
QY TTAGCGGAGAGATTTCTTGAGAAAGTTGAAAGATGCGCAAAATGCTTGTGAAAGGAGAG 1380
Db TTAGCGGAGAGATTTCTTGAGAAAGTTGAAAGATGCGCAAAATGCTTGTGAAAGGAGAG 1380
QY CCTGCTGGAAGATTAACAACAAAGATGATTAATCGTTTACATTTGCTGGAGGAGGAGCACA 1440
Db CCTGCTGGAAGATTAACAACAAAGATGATTAATCGTTTACATTTGCTGGAGGAGGAGCACA 1440
QY TAACACTAAAGAGGCTGAAAGGCTCGTACTTTTGGTCAAACAAGATTCCTGGAATGCTAG 1500
Db TAACTAAAGAGGCTGAAAGGCTCGTACTTTTGGTCAAACAAGATTCCTGGAATGCTAG 1500
QY ATGTTTGAAGAAACAAATTTGCTGATGGAAGGATCAAGAGAGGAGGAGGATTTCAAT 1560
Db ATGTTTGAAGAAACAAATTTGCTGATGGAAGGATCAAGAGAGGAGGAGGATTTCAAT 1560
QY ATGATGCTGTTAGCTTTCGTAATGCTGATGATATATATGTTTGGTCAAACAAGATTCCT 1620
Db ATGATGCTGTTAGCTTTCGTAATGCTGATGATATATATGTTTGGTCAAACAAGATTCCT 1620
QY AGCAAGAGAGAGAGAGAGAGAGGATCAAGAGAGGATCAAGAGAGGATCAAGAGAGGAG 1680
Db AGCAAGAGAGAGAGAGAGAGGATCAAGAGAGGATCAAGAGAGGATCAAGAGAGGAG 1680
QY AGSAGCTTTGTAAGGATCAAAACCCAGTCCATCCAAAGGCTTAAAGTCAAGAGAGGAG 1740
Db AGSAGCTTTGTAAGGATCAAAACCCAGTCCATCCAAAGGCTTAAAGTCAAGAGAGGAG 1740
QY CTAGGCACTCTATGCTGCTGCAACCGAGGCTATCTCTGCTGAGGAGGAGGAGGAGGAG 1800
Db CTAGGCACTCTATGCTGCTGCAACCGAGGCTATCTCTGCTGAGGAGGAGGAGGAGGAG 1800
QY CCCGAGAGAGTGAATATCTGCAATCAAAAGTCTGTTGCTGCTGCAAGAGAGGAGGAG 1860
Db CCCGAGAGAGTGAATATCTGCAATCAAAAGTCTGTTGCTGCTGCAAGAGAGGAGGAG 1860
QY TCGGCACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db TCGGCACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
QY AGTTGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db AGTTGCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY AGATCAACAGGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db AGATCAACAGGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
QY AGGATGATGAGAGAGGAGTCCGAGAGATTTAGAGAGTGAATCCCAAGAGTCCGATGA 2100
Db AGGATGATGAGAGAGGAGTCCGAGAGATTTAGAGAGTGAATCCCAAGAGTCCGATGA 2100
QY CTGTTACGGCTCCATGAGAGATGCAATGATGCTCCCTCTCTGCTGCTGCTGCTGCTGCT 2160
Db CTGTTACGGCTCCATGAGAGATGCAATGATGCTCCCTCTCTGCTGCTGCTGCTGCTGCT 2160
QY CTCCAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db CTCCAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
QY CTTTGAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db CTTTGAAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280

Db 983 TGCAGAGCTTCAAGATTTGGCTCTAGCATGCTTGAATCTCATGGATACAC 1042
Qy 988 CACTTGAAGAGCAGAGATGTTTCAAGATTAACATGCAATATGCTGCTTCAAGACAG 1047
Db 1043 CACTTGAAGAGCAGAGATGTTTCAAGATTAACATGCAATATGCTGCTTCAAGACATG 1102
Qy 1048 AGATACTGAACCAACACCCCTCTCCACAGATTTCTGAATATATGTCGAATCTGAGGTGT 1107
Db 1103 AATAAATGAGCTTAACAGCTCTCTACTGACTTCTCAGCTACGTGGAATCTGAAGTTT 1162
Qy 1108 TAAGGCTTGAACAACTGAAGCAAGTAAAGATGAAAGATCTTGTTTTAAAAAGAAAGCAG 1167
Db 1163 TAAGACTTGAACAGCTTAAGCGAGCAAGATGAAGACCTTGTCTTAAAAAGAAAGACAG 1222
Qy 1168 AACTAGAAGCAGATGAAGACGCTCATCTTGTGTCGAGAGGAGTATGCAAGGACT 1227
Db 1223 AACTGGAAGATCATAGGAGCGTCTCATTTGATTGCGGAGGAGGCTATGCACTGAAT 1282
Qy 1228 TTAGCAATTGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTTGAACAAATG 1287
Db 1283 TTAGCGATGAGGCTATTGAGCGAGGAGCTTGTGATCCTTCGCTGCTGGAACAAATG 1342
Qy 1288 AAGCTCACATTTGCAACAGTGAAGAGAAAGCTTTTAGCCGGAAGAGATATCTTTGAAAG 1347
Db 1343 AGGCTCACATTTGCCAGTGAAGAGAGAGGCTTTTAGCAGGAAGGATATCTCGAAGAG 1402
Qy 1348 TTGAAGATGCGAAATGCTTGTGAAGAGAGAGGCTTGGTGGAGAGATTAACAAAGATG 1407
Db 1403 TTGAAGATGCTTGAATGATGTTGAGGAGGAGCTTGGTGGAGAGATTAACAAAGATG 1462
Qy 1408 ATAATCTTACAATGCTGGAGGGAGCAGCATCTAACACTAAAGAGGCTGAAAGGCTC 1467
Db 1463 ACAATCGTTATAATGCGGAGGGGGGCCCATTTGACATCAAAAGGCTGAGAAAGCTC 1522
Qy 1468 GTACTTTGGTCAACAAAGATTTCTGGAATGGTAGATGTTTGAAGCAAAAAATTTGCTGCAT 1527
Db 1523 GTATTTTGGTTAACAAGATCCAGGACTGGTAGATGTTTAAACCACAAAAAATTTGAGCCT 1582
Qy 1528 GGAAGAAATGAACAGAGAGAGGAGGATTTACATATGATGTTGTTAGCTTTCGTCATGC 1587
Db 1583 GCGAGGCGAGAAAGAGGAA---AGGAAATTCATACATACGACGCTGTCGCGCTTATGTCATGC 1639
Qy 1588 TTGATGAATATATGTTGCTGTCAGGAGAAAGCAAGCAAGAGAGAGCAAGAGGATC 1647
Db 1640 TTGAAGAGTACATGATGTTTGGCGAGAGAGAGCTAACAAGAGAGGAGCAAGGATC 1699
Qy 1648 AGAAGAGCTCCAGGATCAGCTCAAAGCGGAGGAGGAGCTTTGTAGCGATCAAAACCCCA 1707
Db 1700 AGAAGAAATCCAGGATCACTCAAGCTGAGCAGGAAGGCTCTATGGATCAAAACCAA 1759
Qy 1708 GTCATCAAGCCCTTAAGTACAAAGAGGACCTTAGGCACTCTATGGTGGTGGCAAC 1767
Db 1760 GTCCTTCAAGCTCAAAGCAGCAAGAGGCGCTTAGGCACTCCATGGGTGGTGGCAAC 1819
Qy 1768 GAAGGCTATCTTGGTGGAGCCACCATGCAACCCCGAGAGCTGATATACTGATTTCA 1827
Db 1820 GAAGACTGTCTTGGTGGAGGCACTATGCAAGCCCCCAAAACAGATATACTGCACTCCA 1879
Qy 1828 AGTCTGTCTGTGCTGCAAGAAACTGAAGAAATCGGCACTTTGTCCCTTAGTAGTAGTA 1887
Db 1880 AGACAGCTCTGTCGCGCCCAAGAGGCTGAAGATTTGGGCGCTTTATCTCC---TAGTAGTA 1936
Qy 1888 GAGGTTTGAATGTCGCGGATGCTCTATCAAGAGGTTGCTTTTCAATGCCAGTACTCTAC 1947
Db 1937 GAGGCTTGAACATTTGCGCGCTCTCCCATCAAGAGTTATCTTTCAAGCGCAAGCACTTTGC 1996
Qy 1948 GTGAGCGGAGACACTCGTAAACCTTTTGTCTGATGCACACCGAGGAAACAGTGTCTCGT 2007
Db 1997 GAGAGGAGAGAAACACCGCGAAGCTTTTGGCCAGATCATGCCAGGAAACAAAGTCTCAC 2056
Qy 2008 CGAGCGCTGTGCGCCCTTATACCAATAAACAATGAGGATGATGAGAACAGGACTCCGAAGA 2067

Db 2057 CGATGCTTACGCGGCCCATCTCCAGCGCCACAGAGGAAGAGAACAAAACCCCGAGCAT 2116
Qy 2068 CATTTACAGCACTGAATCCCAAGACTCCGATGACTGTTTACGCTCCAATGCAATGGCAA 2127
Db 2117 TTGTAGGAGGCTCAATCGAAAGCGCGGACAGTACGCTCTCTATGCAATGGCGA 2176
Qy 2128 TGATCCCTCTCTGCGCCAAAGGTTTCAGCAATCCAGTTCCTTCTTGTACGACACAGC 2187
Db 2177 TCACACCAAGCGCTGGCTAACCAAGGTCATAGCCACTCTCTGCAACCCCTTTTCCAGGAGAG 2236
Qy 2188 CAGA---GGTAACATTGCGAGGAGGACATCGACTACTCTTTGGAAGAAAGGCGGCTCGCA 2244
Db 2237 CAGATTCGCCAGACTGCGAGCGGACATCGAGTACTCGTTGGAAGAGGCGGCTCGCTG 2296
Qy 2245 TCTATCTGCGCCAGCAAAATGGTTTAACTGTTGATCAATTTATGTAGTGTGAAATCTG 2304
Db 2297 TTTACTTGGCCAGCAAGTGGCTTAA-ACTCGTCAATTCACGTAGTTGAAAGCTGAACCTG 2355
Qy 2305 ACTGCATTTTCTTGTGCGGTGGCCATTG 2331
Db 2356 ACTGCGATTTCTCGTCAGAGGCCAATG 2382

RESULT 4

US-10-619-685-3
; Sequence 3, Application US/10619685
; Publication No. US20040096875A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Resources, Ministry of Agriculture,
; APPLICANT: Forestry and Fisheries
; APPLICANT: Japan Science and Technology Corporation
; TITLE OF INVENTION: A NOVEL GENE FOR CONTROLLING LEAF SHAPES
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/619,685
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US/09/667,475D
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 4574
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-619-685-3

Query Match 41.4%; Score 1021.2; DB 16; Length 4574;
Best Local Similarity 69.6%; Pred. No. 9.1e-278;
Matches 1974; Conservative 0; Mismatches 3; Indels 861; Gaps 9;

Qy 469 TTAGACAGTCAAAATCAGAAAGCTTTCATGGCTTAAGAGAGAGTGAATGCAATTTGTTCCGT 528
Db 1177 TTTTGCAGTCAAAATCAGAAAGCTTTCATGGCTTAAGAGAGAGTGAATGCAATTTGTTCCGT 1236
Qy 529 ATTTGGAAGAAATGAAAAAGAAAGGTCGAACGATGGAACCACTTTGTTCAATGTCATAG 588
Db 1237 ATTTGGAAGAAATGAAAAAGAAAGGTCGAACGATGGAACCACTTTGTTCAATGTCATAG 1296
Qy 589 AGCAGATTAAAGAAATTTCTGCTGAAATAAGGCCAGCCGATTTGTTCCCTTTAAAGTTTC 648
Db 1297 AGCAGATTAAAGAAATTTCTGCTGAAATAAGGCCAGCCGATTTGTTCCCTTTAAAGTTTC 1356
Qy 649 CGGTTGATCAGTGTGACCTGTCTATTAAAGAAAGCTTGAATGAGTTGACGAAGGACTCGAAT 708
Db 1357 CGGTTGATCAGTGTGACCTGTCTATTAAAGAAAGCTTGAATGAGTTGACGAAGGACTCGAAT 1416
Qy 709 CCCTTCAGAAAGAGA----- 723
Db 1417 CCCTTCAGAAAGGAGAGGTCATCATCAATAACCATCTTTATCCATTTTTCACGACTCAT 1476
Qy 724 ----- 723
Db 1477 GTTGTGATCGTGTCTCTATCTATCAAGAAATCTCTTTTCATTTCTTGATATAAAATCTCACTA 1536

QY 724 -----AGCGGATCGGCTAAAGCAAGTGTAGTAAACATTG 758
DB TGCATATACATGTTTGTCTCAGACGCGATCGGCTAAAGCAAGTGTAGTAAACATTG 1596
QY 759 AATCTTTGCAATCTTATGTGAGGTGCTTGGCATAGATTTTCAAGCAACAGTATATGAG 818
DB AATCTTTGCAATCTTATGTGAGGTGCTTGGCATAGATTTTCAAGCAACAGTATATGAG 1656
QY 819 GTGACCCCTAGCTTGGACGAAGCTGAAGATCAAGAACTGAGCAACATCAATTTGAG 878
DB GTGACCCCTAGCTTGGACGAAGCTGAAGATCAAGAACTGAGCAACATCAATTTGAG 1716
QY 879 AGGCTTGCTGCTGCCCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAA----- 933
DB AGGCTTGCTGCTGCCCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAAAGGTC 1776
QY 934 ----- 933
DB 1777 AGCAATGCTGTACCAATTTAGAGGTATCAATGAACACATTTTCAGTCTTTAACTTGGTTAA 1836
QY 934 -----AGCTTCAAGATTTTGTCTTAGCATGCTCGAGTATGGATCTCATGG 981
DB TCTGATTTCTGGCAGCTTTCAAGATTTTGTCTTAGCATGCTCGAGTATGGATCTCATGG 1896
QY 982 ATACTCCACTTGAAGAGCAGCAGATGTTTTCAGAAATATAACATGCAATATGCTGCTTCAG 1041
DB ATACTCCACTTGAAGAGCAGCAGATGTTTTCAGAAATATAACATGCAATATGCTGCTTCAG 1956
QY 1042 AACAGAGATAACTGAACCAACACACCTCTCCACAGATTTTCTGAAATAT----- 1091
DB 1957 AACAGAGATAACTGAACCAACACACCTCTCCACAGATTTTCTGAAATATGTAATTTATC 2016
QY 1092 ----- 1091
DB 2017 ATCACTGAGATTGCAAAATTTATGTTGCTATGTTGTTATATTTTCAATTAAGATATGAAT 2076
QY 1092 -----GTCGAATCTGAGGTGTTAAGGCTTGAACAACTG 1124
DB 2077 GTTTCATCGACTACTTATACTGTAGTCTGAAATCTGAGGTGTTAAGGCTTGAACAACTG 2136
QY 1125 AAAGCAAGTAAAGATCTGTTTAAAAAAGAAAGCAAGCTAGAGAGCATAGA 1184
DB 2137 AAAGCAAGTAAAGATCTGTTTAAAAAAGAAAGCAAGCTAGAGAGCATAGA 2196
QY 1185 AGACGTGCTCATCTTGTGGCAGAGAGGTATGCAAGAGGAGTTAGCATTTGAAGCTATT 1244
DB 2197 AGACGTGCTCATCTTGTGGCAGAGAGGTATGCAAGAGGAGTTAGCATTTGAAGCTATT 2256
QY 1245 GAAGCT----- 1250
DB 2257 GAAGCTGGTAAAGATCTCTCTGCTTACTGCTGCTTTTATGTCCTGACAAAGTCATACCA 2316
QY 1251 -----GGAGCTATTGATCCCTCATAGT 1273
DB 2317 GACAGAGTTTATATCTGCTGCTGTTCTGTTTCGAGGAGCTATTGATCCCTCATAGT 2376
QY 1274 ACTTGAACAAATTTGAAGCTCATTTGCAACAGTGAAGAGGAGCTTTTAGCCGAAGGA 1333
DB 2377 ACTTGAACAAATTTGAAGCTCATTTGCAACAGTGAAGAGGAGCTTTTAGCCGAAGGA 2436
QY 1334 TATCTTTGAGAAAGTTGAAGATGCAAAATGCTTTGGAAGAGGAGGCTGCTGGAAGA 1393
DB 2437 TATCTTTGAGAAAGTTGAAGATGCAAAATGCTTTGGAAGAGGAGGCTGCTGGAAGA 2496
QY 1394 TTACAACAAA----- 1403
DB 2497 TTACAACAAAGTATGAGTCTAGCTGAAGCTACGTTGTTTGTATATTTGTTAGCAAA 2556
QY 1404 -----GATGATAATCGTTTACAATGC 1423
DB 2557 TAATGTGTTACTGATATCTCTGGCTTTGGCTTTTATAGGATGATATCGTTTACAATGC 2616
QY 1424 TGGAGGGGAGCACATCTTAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTGGTCAACAA 1483

DB 2617 TGGAGGGGAGCACATCTTAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTGGTCAACAA 2676
QY 1484 GATTCCT----- 1490
DB 2677 GATTCCTGGTAAATGTTACTCAATGATTTATGTTTGGAACTTCTCCTTATCAAGTGCATAT 2736
QY 1491 ----- 1490
DB 2737 TTAATTTACAATTTTAACTCTTGGCAATTAACAATCTGATATCTCTGCTGATTTGTGCTG 2796
QY 1491 -----GGAATGTTAGATGTTTGGAGAACAAAAATTTGCTCATGGAATAATGAACAGGAAA 1546
DB 2797 AGCAGGAATGTTAGATGTTTGGAGAACAAAAATTTGCTCATGGAATAATGAACAGGAAA 2856
QY 1547 GGAGGAATTTACATATGATGCT----- 1569
DB 2857 GGAGGAATTTACATATGATGCTGTTTCTTACTCTTACACATTTACATTTGATCGGCT 2916
QY 1569 -----GTTAGCCTTTCTGCA 1583
DB 2917 CTATTTTGTGTTCTTGTGTAAGTGCCTTTCTTGCATTTCTTACAGGTTTAGCCTTTCTGCA 2976
QY 1584 ATGCTTGTATGATATATGTTTCTGTCAGAGAAAGAGCAAGAGAAAGAGCAACAAAG- 1642
DB 2977 ATGCTTGTATGATATATGTTTCTGTCAGAGAAAGAGCAAGAGAAAGAGCAACAAAG 3036
QY 1643 ----- 1642
DB 3037 GTATTTATGCTCTCGCCTAATATTCATGTTATTTCTAAATCATCTTTTACCTTCTGTGAA 3096
QY 1643 -----GGATCAGAAGAGCTCCAGATCAGATCAGCTCAAAG 1674
DB 3097 TACGCTCTAATCTTGAATATATCCTCGAGGATCAGAAGAGCTCCAGATCAGCTCAAAG 3156
QY 1675 CGGAGCAGGAAGCTTTGTAGCGGATCAAAACCCAGTCCATCCAGCCCTCAAGTCAAGAA 1734
DB 3157 CGGAGCAGGAAGCTTTGTAGCGGATCAAAACCCAGTCCATCCAGCCCTCAAGTCAAGAA 3216
QY 1735 AGGCACCTAGGACCTCTATGCTGGTGGTGCNAACCGAAGCTATCTCTTGGTGGAGCCACCA 1794
DB 3217 AGGCACCTAGGACCTCTATGCTGGTGGTGCNAACCGAAGCTATCTCTTGGTGGAGCCACCA 3276
QY 1795 TGCAACCCCGAAGACTGATATCTGCAATCAAAAGTCTGTTGCTGCTGCCAAGAAAACTG 1854
DB 3277 TGCAACCCCGAAGACTGATATCTGCAATCAAAAGTCTGTTGCTGCTGCCAAGAAAACTG 3336
QY 1855 AAGAAATCGGCACCTTTGTCCTCTAGTA----- 1881
DB 3337 AAGAAATCGGCACCTTTGTCCTCTAGTAAGCCCTACTAGCTATCATGTGTCGATATATTTTC 3396
QY 1882 -----GTAGTA 1887
DB 3397 TTTTTCCTCTTATTTTCACTTGAAACATATGTTAACTCAAGAAACAAATATCAAGGTAGTA 3456
QY 1888 GAGGTTTGGACATTTGCCGGATTTGCTATCAAGAAGTTGTCTTTTCAATGCCAGTACTCTAC 1947
DB 3457 GAGGTTTGGACATTTGCCGGATTTGCTATCAAGAAGTTGTCTTTTCAATGCCAGTACTCTAC 3516
QY 1948 GTGAGACGGAGACACCTCGTAAACCTTTTGTCTAGATCAACCAAGGAAAACTGCTCGT 2007
DB 3517 GTGAGACGGAGACACCTCGTAAACCTTTTGTCTAGATCAACCAAGGAAAACTGCTCGT 3576
QY 2008 CGACGCTGTCGCGCTTATCAACCAATTAACCTGAGGATGATGAGAACAGGACTCCGAGAA 2067
DB 3577 CGACGCTGTCGCGCTTATCAACCAATTAACCTGAGGATGATGAGAACAGGACTCCGAGAA 3636
QY 2068 CATTTACAGCACTGAATCCCAAGACTCCGATGACTGTTAGCGCTCCCAATGAGATGCAAA 2127
DB 3637 CATTTACAGCACTGAATCCCAAGACTCCGATGACTGTTAGCGCTCCCAATGAGATGCAAA 3696
QY 2128 TGACTCCCTCTCTGGCCAAACAAAGTTTCAGCAACTCAGTTTCCCTTGTGTTTACGCAAGC 2187

Db 3697 TGACTCCCTCTCTGCGCCAAAGAGTTTTCAGCAACTCCAGTTCCTTGTGTTACGCAAGC 3756
Qy 2188 CAGAGGTACATTGCGAGGAGCATCGACTACTCTCTTTCAGAAAGCGGCTCGCCATCT 2247
Db 3757 CAGAGGTACATTGCGAGGAGCATCGACTACTCTCTTTCAGAAAGCGGCTCGCCATCT 3816
Qy 2248 ATCTGCGCCAGGCAAAATGGTTTAACTGTTGATCAATTTATGTAGTGTGAAATCTGACT 2307
Db 3817 ATCTGCGCCAGGCAAAATGGTTTAACTGTTGATCAATTTATGTAGTGTGAAATCTGACT 3876
Qy 2308 GCATTTCTTGTGCGTGGCCATGCGTATGTTGGTCAAAAGTGGCTTCCAGTAG 2367
Db 3877 GCATTTCTTGTGCGTGGCCATGCGTATGTTGGTCAAAAGTGGCTTCCAGTAG 3936
Qy 2368 CACTATTCTGATTCTAGTCAATTTGTTTAAATGTTTCTACACCAAGTAAACAGCTCTAT 2427
Db 3937 CACTATTCTGATTCTAGTCAATTTGTTTAAATGTTTCTACACCAAGTAAACAGCTCTAT 3996
Qy 2428 ACATTAGCTTGTCTACTA 2445
Db 3997 ACATTAGCTTGTCTACTA 4014

RESULT 5
US-10-437-963-95930/c
; Sequence 95930, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 95930
; LENGTH: 784
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_94075C.1
US-10-437-963-95930

Query Match 27.3%; Score 674.4; DB 17; Length 784;
Best Local Similarity 97.7%; Pred. No. 6e-180;
Matches 717; Conservative 0; Mismatches 11; Indels 6; Gaps 3;
Qy 1714 CCAAGCCCTTAAGTACAAAGAGCCACTAGGCACTCTATGGTGGTGCAGAAACCGAAGGC 1773
Db 778 CCAAGCCCTTAAGTACAGGAGGACC-AGGCACTCTATGGTGGTGCAGAAACCGAAGGC 720
Qy 1774 TATCTCTTGGTGGAGCCACCATGCAACCCCGAGACTGATATCTGCAATTCAGAGTCTG 1833
Db 719 TATCTCTTGGTGGAGCCACCATGCAACCCCGAGACTGATATCTGCAATTCAGAGTCTG 660
Qy 1834 TTGCTGCTGCCAGAAACTGAGAAATCGGCACTTTGCTCCCTAGTAGTAGAGGTT 1893
Db 659 TTGCTGCTGCCAGAAACTGAGAAATCGGCACTTTGCTCCCTAGTAGTAGAGGTT 603
Qy 1894 TGGACATTGCGGATTCGCTTATCAAGAGTGTCTTTCAATGCGAGTACTCTACGTGAGA 1953
Db 602 TGGACATTGCGGATTCGCTTATCAAGAGTGTCTTTCAATGCGAGTACTCTACGTGAGA 543
Qy 1954 CGGAGACACCTCGTAAACCTTTTGTCTCAGATCACACAGGAAACAGTGTCTCGTGCAGC 2013
Db 542 CGGAGACACCTCGTAAACCTTTTGTCTCAGATCACACAGGAAACAGTGTCTCGTGCAGC 483

Qy 2014 CTGTGCGCCCTATCAACCAATTAACACTGAGGATGATGAGAAAGGACTCCGAGACATTTA 2073
Db 482 CTGTGCGCCCTATCAACCAATTAACACTGAGGATGATGAGAAAGGACTCCGAGACATTTA 423
Qy 2074 CAGCACTGAATCCCAAGACTCCGATGACTGTTACGGCTCCCAATGCGATGGCAATGACTC 2133
Db 422 CAGCACTGAATCCCAAGACTCCGATGACTGTTACGGCTCCCAATGCGATGGCAATGACTC 363
Qy 2134 CCTCTCTGCGCCAAAGGTTTTCAGCAACTCCAGTTCCTTGTGTTACGACAGCCAGAGG 2193
Db 362 CCTCTCTGCGCCAAAGGTTTTCAGCAACTCCAGTTCCTTGTGTTACGACAGCCAGAGG 303
Qy 2194 TAAACATTCAGAGAGACATCGACTACTCTTTCAGAAAGGCGGCTCGCCATCTATCTGG 2253
Db 302 TAAACATTCAGAGAGACATCGACTACTCTTTCAGAAAGGCGGCTCGCCATCTATCTGG 243
Qy 2254 CCAGGCAATGGTTTAACTGTTGATCAATTTATGTAGTGTGTTGAAATCTGACTGCAATTT 2313
Db 242 CCAGGCAATGGTTTAACTGTTGATCAATTTATGTAGTGTGTTGAAATCTGACTGCAATTT 183
Qy 2314 TCTTCTCGTGGCCATTCGCTATGTTGGTCAACATAGTTCGGCTTTCCAGTAGCACTAT 2373
Db 182 TTTTGTGCGTGGCCATTCGCTATGTTGGTCAACATAGTTCGGCTTTCCAGTAGCACTAT 123
Qy 2374 TCTGATTTACTGCAATGTTTAAATGTTT--CTACAAACCAAGTAAACAGCTCTATACAT 2431
Db 122 TTTGATTTACTGCAATGTTTAAATGTTTCTACAAACCAAGTAAACAGCTCTATACAT 63
Qy 2432 TAGCTTGTCTACTA 2445
Db 62 TAGCTTGTCTACTA 49

RESULT 6
US-10-437-963-81367
; Sequence 81367, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 81367
; LENGTH: 709
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(709)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_80900C.1
US-10-437-963-81367

Query Match 24.8%; Score 611.6; DB 17; Length 709;
Best Local Similarity 98.1%; Pred. No. 3.6e-162;
Matches 629; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 4 AAAATATTTCAATACACTACACTCTCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63
Db 71 AAAATATTTCAATACACTACACTCTCGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 130

QY 64 CTCCTCCGCTCTCTCGCATCTGAGGCTCCGATCGCCGCGACCCAGCAGCAATCCGCC 123
Db 131 CTCCTCCGCTCTCTCGCATCTGAGGCTCCGATCGCCGCGACCCAGCAGCAATCCGCC 190
QY 124 GCCCGCTCTCGGCTCCCGCTCGACGAGACCGCGCGCGGCGAGGGCTAGTGT 183
Db 191 GCCCGCTCTCGGCTCCCGCTCGACGAGACCGCGCGCGGCGAGGGCTAGTGT 250
QY 184 CTTTCGACCTTCGCGATGAGTACGGGTGAAGGACACAGCTTCACAGATGTCGACGACAT 243
Db 251 CTTTCGACCTTCGCGATGAGTACGGGTGAAGGACACAGCTTCACAGATGTCGACGACAT 310
QY 244 GCGATTTCGCTTCTACTGGAGCTCAATGTGATTTGGGATGAGTTCGCTGAGCCCGACAGA 303
Db 311 GCGATTTCGCTTCTACTGGAGCTCAATGTGATTTGGGATGAGTTCGCTGAGCCCGACAGA 370
QY 304 CAGGGACAGGATGCTCTGAGAGCTCGACGAGAGTGCCTGAGGTCTACAGGCGGAAGG 363
Db 371 CAGGGACAGGATGCTCTGAGAGCTCGACGAGAGTGCCTGAGGTCTACAGGCGGAAGG 430
QY 364 TGACACAGCGAAACCGGAGCCGCCAGCTCGGAAGGCCATCGCGAGGCGGAGGCGAG 423
Db 431 TGACACAGCGAAACCGGAGCCGCCAGCTCGGAAGGCCATCGCGAGGCGGAGGCGAG 490
QY 424 AGCTCCCGGCTCTCTCTCAGCCATGGCGAGCCGCCCGCTGCACGTCTAGACAGTCAAAATC 483
Db 491 AGCTCCCGGCTCTCTCTCAGCCATGGCGAGCCGCCCGCTGCACGTCTAGACAGTCAAAATC 550
QY 484 AGAAGCTTCATGCTTAAGAGAGGAGTTGAATGCAATTTGTCGATTTTGGAGAAATGA 543
Db 551 AGAAGCTTCATGCTTAAGAGAGGAGTTGAATGCAATTTGTCGATTTTGGAGAAATGG- 609
QY 544 AAAAGAAAAGGTTCGAACCATGGAACCACTTTGTTTCATGTCATAGACAGATTAAGAAAA 603
Db 610 -AAAGAAAAGGTTCGAACCATGGAACCACTTTGTTTCATGTCATAGACAGATTAAGAAAA 668
QY 604 TTTCTGCTGAATAAGGCGAGCCGATTTTGTTCCTTTAAA 644
Db 669 TTTCTGCTGAATAAGGCGAGCCGATTTTGTTCCTTTAA 709

RESULT 7

US-10-437-963-98451
; Sequence 98451, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5321)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 98451
; LENGTH: 2045
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MKT4530_96356C.1
US-10-437-963-98451

Query Match 22.8%; Score 563; DB 17; Length 2045;
Best Local Similarity 76.6%; Pred. No. 4.4e-148;
Matches 689; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
QY 198 ATGAGTAGCGGTGAAGACACAGGTTCCACAGATGTCGACGACATGCGATTTCGTTCTTA 257

Db 1 AGGAGCGCAATCCGAGGAGCGCTCTCTCCGATGAAGACCATATGCGGATCCCTCATG 60
QY 258 CTGGAGCTCAATGTGATTTGGGATGAGGTCCGCTGAGCCCGACACGACGAGGACAGGATG 317
Db 61 CACAGGCTCCAGGTCATCTGGGACGAGGTCTGGGAGCGCCGAGGCGCGGAGGACAGGATG 120
QY 318 CTGCTGGAGCTCGAGCAGGAGTGCCTGGAGGTCTACAGGCGGAAGGTTCGACACAGCGCAAC 377
Db 121 CTGCTGGAGCTCGAGCAGGAGTGCCTGGAGGTCTACAGGAGGAGGTTGGAACAGGCGAAC 180
QY 378 CGGAGCGCGCCCGAGCTCGGGAAGGCCATCGCCAGGCGGAGGAGGAGTCTCGCGGATC 437
Db 181 CGCTCCGCGCCAGCTCGCGCAGCGATCGCTCAATATGAAGCTGAACCTCGCGCCATC 240
QY 438 TGCTCAGCATGCGGAGCGCCCGCTGCACGTTAGACAGTCAATCAGAACTTCATGGC 497
Db 241 TGCTCCGCAATTGCGGAGACGACCGTCCAGTCAGGAGTCTAATCAGAAAGCATGTGGA 300
QY 498 TTAAGAGAGGAGTTGAATGCAATTTGTTCCGTTTGAAGAAATGAAAAAGAAAAGGTC 557
Db 301 TTACGGGATGAGCTCGGTGCANTATTACCATACCTAGAGAGATGAAGAGGAAGGTT 360
QY 558 GAACGATGGAACCACTTTGTTTCATGTATAGAGCAGATTAAGAAAAATTTCTCTGAATA 617
Db 361 GAAAGATGGAACCACTTTGTTTCATGTGTAGGAAGGATAAAGAAAGATCTCATCTGAGATA 420
QY 618 AGGCGAGCGGATTTGTTCCCTTTAAAGTTCCGGTTCATGATGATGAGTCTGATTAAGA 677
Db 421 AGGCGAGCAATTTGAGCCCTTTTAAAGTGTCTGTGGATCAATCTGATCTATCAATAAGA 480
QY 678 AAGCTTGATGATGAGTTCAGAGGAGCTCGAAATCCCTTCAGAGGAGAGAGCGATCGGCTA 737
Db 481 AAGCTTGAGAGTTAAGGTGGAGCTGAGTCCCTTGAAAGGAGAGGAGGAGGTA 540
QY 738 AAGCAAGTATGAGAACTTTGAAATTTCTTTCATTCCTTATGATGAGGTCCTTGGCATAGAT 797
Db 541 AAGCAAGTATGAGAACTTTGAAAGCTTTAAGAGCTTTAATCTTTTATGTAGTACTTGTGTCGAC 600
QY 798 TTCAGCAACACAGTATATGAGTGCACCTGACCTGCTGCAAGCAAGCTGAAGGATCAAGAAC 857
Db 601 TTCAGCAACACATATCTGAAATACACCTAGTCTTGTATGAGCTGAAGGCGCAAGGANT 660
QY 858 CTGAGCAACACTCAATTTGAGAGGCTTTGCTGCTGCCGCAACACAGACTCGCTGAAATGAAG 917
Db 661 ATAAGCACTACTCAATTTGAGATGCTAGCATGGGCGATTCAGAGACTTCGTGAAACAAA 720
QY 918 ATCCAAAGGATGCAAAAGCTTCAAGATTTTGTCTTCTAGCATGCTCGAGCTATGGAATCTC 977
Db 721 ATGCAGAGGATGCAAAAGCTTCAAGATCTTGCATCTACCTTGTCTAGAACTATGGAATCTA 780
QY 978 ATGGATCTCCACTTTGAAGGAGCAGCATGTTTCAGAAATATAACATATTCATTCGCT 1037
Db 781 ATGGATCGCCATTTGAAGAGCAGCAGCATACAGAAATATAACATATTCATTCGCT 840
QY 1038 TCAGAACAGAGATAACTGAACCAACACCCCTCTCCACAGATTTCTGAAATATGTCGA 1096
Db 841 TCAGAGGCTGAATTAACAGAACAGAACACCCCTTTCATTCATTCGCTTCTCAACTATGATGA 899

RESULT 8

US-10-437-963-32360
; Sequence 32360, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

;; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
;; FILE REFERENCE: 38-21(53221)B
;; CURRENT APPLICATION NUMBER: US/10/437,963
;; CURRENT FILING DATE: 2003-05-14
;; NUMBER OF SEQ ID NOS: 204966
;; SEQ ID NO 32360
;; LENGTH: 530
;; TYPE: DNA

;; ORGANISM: Oryza sativa

;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT4530_36577C.1

US-10-437-963-32360

Query Match 20.5%; Score 506.6; DB 17; Length 530;

Best Local Similarity 99.0%; Pred. No. 1.7e-132;

Matches 520; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1163 AGCAGAACTAGAGACATAGAGAGCTGCTCATCTTGTGGCGAGGAAGGTTATGCAGA 1222

DB 6 AGGAGAACTTGAAGACATAGAGAGCTGCTCATCTTGTGGCGAGGAAGGTTATGCAGA 65

QY 1223 GGAGTTTACATTGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTTGAACA 1282

DB 66 GGAGTTTACATTGAAGCTATTGAAGCTGAGCTATTGATCCCTCACTAGTACTTGAACA 125

QY 1283 AATTGAAGCTCATTGCAACAGTGAAGAGGAGCTTTAGCCGGAAGGATATCTTGA 1342

DB 126 AATTGAAGCTCATTGCAACAGTGAAGAGGAGCTTTAGCCGGAAGGATATCTTGA 185

QY 1343 GAAAGTTGAAGAGTGCAGAAATGCTT-GTGAAGAGGAAGCTTGGCTGGAAGATTACAACA 1401

DB 186 GAAAGTTGAAGAGTGCAGAAATGCTTGGGGAAGAGGAGCTTGGCTGGAAGATTACAACA 245

QY 1402 AAGATGATAATCGTTACAATGCTGGAGGGAGGACATCTAACACTAAAGAGGGCTGAAA 1461

DB 246 AAGATGATAATCGTTACAATGCTGGAGGGAGGACATCTAACACTAAAGAGGGCTGAAA 305

QY 1462 AGGCTGCTACTTGGTCAACAGATTCCTGGAATGCTAGATGTTTGAAGAACAAAATTG 1521

DB 306 AGGCTGCTACTTGGTCAACAGATTCCTGGAATGCTAGATGTTTGAAGAACAAAATTG 365

QY 1522 CTCGATGGAAAAATGAACGAGGAAAGGAGGATTCACATATGATGGTGTAGCCTTTCGT 1581

DB 366 CTCGATGGAAAAATGAACGAGGAAAGGAGGATTCACATATGATGGTGTAGCCTTTCGT 425

QY 1582 CAATGCTTGATGATATATGTTCTTTCGTCAGAGAAAGCAAGAGAAAGAGACAAA 1641

DB 426 CAATGCTTGATGATATATGTTCTTTCGTCAGAGAAAGCAAGAGAAAGAGACAAA 485

QY 1642 GGGATCAGAAAGAGCTCCAGGATCAGCTCAAGGGGAGGAGG 1686

DB 486 GGGATCAGAAAGAGCTCCAGGATCAGCTCAAGGGGAGGAGG 530

RESULT 9

US-10-424-599-126041

;; Sequence 126041, Application US/10424599

;; Publication No. US20040031072A1

;; GENERAL INFORMATION:

;; APPLICANT: La Rosa Thomas J

;; APPLICANT: Kovalic David K

;; APPLICANT: Zhou Yihua

;; APPLICANT: Cao Yongwei

;; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

;; FILE REFERENCE: 38-21(53223)B

;; CURRENT APPLICATION NUMBER: US/10/424,599

;; CURRENT FILING DATE: 2003-04-28

;; NUMBER OF SEQ ID NOS: 285684

;; SEQ ID NO 126041

;; LENGTH: 2150

;; TYPE: DNA

;; ORGANISM: Glycine max

;; FEATURE:

;; OTHER INFORMATION: Clone ID: PAT_MRT3847_84823C.1

US-10-424-599-126041

Query Match 19.0%; Score 467.8; DB 16; Length 2150;

Best Local Similarity 58.8%; Pred. No. 4.5e-121;

Matches 887; Conservative 0; Mismatches 607; Indels 15; Gaps 4;

QY 240 ACATCGGATTGCTTCTACTGAGCTCAATGTGATTTGGGATGAGGTGCGTGAGCCGAC 299

DB 139 ACATGTGTTCTTGTGTTAAAAAGCTTCAGGAATATGGATGAGTGTGGTGAGCGAC 198

QY 300 ACGACAGGACAGGATGCTGCTGAGCTCGAGCAGGAGTGCTCGAGGTCTACAGCGG 359

DB 199 GAGCAACAGACACAGATGCTTCTCAGTTAGACAGGAGTGCTTGGATGTGTACAAGAG 258

QY 360 AAGTTCGACAGCGCAACCGGAGCGCCCGACGCTCGGGAAGGCCATCGCCGAGCGGAG 419

DB 259 AAGTTCGACAGCGCTGCAAAAGTCAAGGGCGGAGCTACTTCAAGCTCTGTCTGATGTAAG 318

QY 420 GCAGAGCTCGCGGCGATCTGCTCAGCCATGGCGGAGCGCCCGTGCACGTTAGACAGTCA 479

DB 319 CTTGAGCTTTCACCTTCTATCAGCACTTGGAGAAAGAGCTT---TGCTGGAATTCCT 375

QY 480 AATCAGAGCTTTCATGGCTTAAAGAGAGGAGTTGAATGCAATTTGTTCCGTTATTTGGAAGAA 539

DB 376 GAGATACTTCTGGAACTATCAAGAAACAGCTTGCAGCTTATAGCACCACTACTTGAACAG 435

QY 540 ATGAAAAAGAAAAGGTGCAACGATGGAACCACTTTGTTTCATGTCATAGACGAGATTAAG 599

DB 436 TTATGCAACAAAAGGAAGAAAGAAATTAAGGAGTTCTCGGATGTACAGTCACAGATCCAA 495

QY 600 AAAATTTGCTCTGAAATTAAGGCGCAGCG---ATTGTTTCCCTTTAAAGTTCGCGTTGAT 656

DB 496 CAATATGTTGGAGAGATAGCTCGAACTTGAACCTTAATGATGTTTCCACCTGCAGTTGAT 555

QY 657 CAGTCTGACCTGCTCAATTAAGAAAGCTTGATGAGTTGACAAAGACCTCGAAATCCCTTTCAG 716

DB 556 GAGTCTGACCTGCTCCCTGAAAGAGTTGGATGAATATCAATCTGAGCTCCAAGAACTTCAA 615

QY 717 AAGGAGAGAGCGATCGGCTAAAGCAAGTATGATGAGTTGACAAAGACCTCGAAATCCCTTTC 776

DB 616 AAGGAAAAGAGTGAGAGGTTGCACAAGGTTCTTGAATTTGTGAGTACTGTGATGATCTA 675

QY 777 TGTGAGTCTTGGCATAGATTTCAAGCAAAACAGATATATGAGTGCACCTAGCTTTGGAC 836

DB 676 TGTGCTGCTTGGTATGAGACTTCTTCACTACTGCACTGAGGTTTCACTCAAGTCTAAAT 735

QY 837 GAAGCTGGAAG-----ATCAAGAAACCTGAGCAACACTACAAATTGAGAGGCTTGTGCT 890

DB 736 GACTCTACTGTTTCAATCCCAAGAGCATAAGTAAATGACACCTTCAAGAGCTGGCTAAG 795

QY 891 GCGCAAAACAGACTGGTGAATTAAGATCCAAAGGATGCAAAAGCTTCAAGATTTTGGCT 950

DB 796 ACCGCTTAAACGCTGAAAGAGATTAAGAAACAGAGGCTGCACAAAGTCCAAAGATTAGCT 855

QY 951 TCTAGCATGCTGAGCTATGGAATCTCATGGTACTTCCACTTGAAGAGCAGCAGATGTTT 1010

DB 856 TCTCAGTTGATGATCTTTGGAACTTAATGTCAGCTTCTGTTGATGAAGTCACTGTTCTGGTGGCCTT 915

QY 1011 CAGAATATAACATGCAATATTTGCTGCTTCAGAACAAAGAGATAAATCTGAACCAACACCTC 1070

DB 916 GACCATGTTACCTGTAATATGTCAGCTTCTGTTGATGAAGTCACTGTTCTGCTGGTGGCCTT 975

QY 1071 TCACAGATTTCTCTGAATTTATGTCGAATCTGAGGTGTTAAGGTTTGAACCACTTGAAGCA 1130

DB 976 GCTCTGGATCTGATTCAGCAGGCTGAAAGTGAAGTTGAGAGCTTTCATCAGCTGAAAGCC 1035

QY 1131 AGTAAGATGAAGATCTTCTTTTAAAGAAAGAGCAGAACTAGAAAGAGCATAGAAGACGT 1190

DB 1036 AGCAGGATGAAGGAAATTTGCTTTTCAAGAAAGCAAGAGAGCTCGAAGAGATATTGTCCTT 1095

Db 1089 CTCTGAGGCCAAAGATGAGGAGTTCAGCAGGAAGGATGTTCTTGAGAGATGAGGAAATG 1148
Qy 1358 GCAAAATGCTTGTGAAGAGAGCCCTGGCTGGAAGATTACAAAGATGATAATCGTTA 1417
Db 1149 GCAGGCTGCACCTAGAAAGAGGAATCTTGGCTTGAGGAGTACAACAGAAATGAGAACAGATA 1208
Qy 1418 CAATGCTGGAGGGAGGACACATCTAAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTGGT 1477
Db 1209 CAATGCTGGCAAGGACATCTCTTGTCTGAAGCGTCCGAGAAAGCGCTGCTTGGT 1268
Qy 1478 CAACAAGATCTCTGGAATCGTATGTTTTTGAGAACAAATAATGCTGCAATGGAATAATGA 1537
Db 1269 CAGCAAAATGCCAGCAATGGCAGAGCCCTTGATCAAAAGGTAATTTGCTTGGGAGAGGA 1328
Qy 1538 ACAGGAAGAGGAGATTTCACATATGATGCTGTAGCCCTTTCGTCATCTTGTATGAATA 1597
Db 1329 GAGGGG---TGCNAATTTGAGTACGATGGTGATGCTCTCGGACATGCTGGAGGAT 1385
Qy 1598 TATGTTTCGTTCTGAGGAGAAAGCAAGAGAGAGAGAGACAAAGGATCAGAAAGCT 1657
Db 1386 CAACAATACAGGAAGAT 1445
Qy 1658 CCAGGATCAGCTCAAGCGGAG 1679
Db 1446 GCTGGGTCAAGGACAGGGGAG 1467

RESULT 11

US-10-437-963-28884/c
; Sequence 28884, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 28884
; LENGTH: 2411
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_33439C.1
US-10-437-963-28884

Query Match 15.8%; Score 389.8; DB 17; Length 2411;
Best Local Similarity 55.94; Pred. No. 6.3e-99;
Matches 852; Conservative 0; Mismatches 647; Indels 24; Gaps 5;
Qy 240 ACATGCGATTCTCTACTGAGCTCAATGTGATTTGGGATGAGGTGCGTGAGCCCGAC 239
Db 2145 ACCTGTGGATCTTCTGCTGCAGAACTGCAGTTGATGGGATGAGTGGTGAGAGCGAC 2086
Qy 300 ACAGGAGGACAGGATGCTGTGGAGCTCGAGCAGGAGTCCCTGGAGGTCTACAGCGG 359
Db 2085 GAGGACCGCACAAAGTCTCTGTACCACTAGATCAGGAGTGCCTGGATGTTTCAAGAGG 2026
Qy 360 AAGTGCAGCAGGACCGGACCGCGCCAGCTGCGGAGGCCATCGCGAGGGCGAG 419
Db 2025 AAAGTTGACAGGCCACCGACTTAGGGATCTCTTATCCAGGCGCTGGATGACTCGAAG 1966
Qy 420 GCAGAGCTCCCGGCATCTGCTAGCCATGGGGAGCGCCGCTGACAGCTTAGACAGTCA 479
Db 1965 ATTGAGCTGCCAGACTTTTATCTGCTTAGGAGAGAGAGGCCAT---AGCAAGACTCCT 1909

Qy 480 AATCAGAAAGCTTTCATGGCTTAAGAGAGGAGTGAATGCAATTTGTTCCGTATTTTGAAGAA 539
Db 1908 GAGAAGAGCTCTGGGACAATCAAGCAACAACTTGTCTGTATAGCCCAACACTCGAGAAG 1849
Qy 540 ATGAAAAAAGAGGTGCAACGATGGAACAGTTTGTTCATGTCTATAGAGCAGATTAAG 599
Db 1848 TTGAACAAGCAGAAAAACGAAAGAGTGGGGAATTTGTAAATGTTCACTCAAAATTTGAT 1789
Qy 600 AAAATTTCTGCTGAATAAAGCCAGCCGATTTTGTTCCTTTTAAAGTTCC-----G 650
Db 1788 CAATATATGTTGATGATTTGCTGGCACTACGAGGTTGGGTGAGAAGTGGCAACGCCACAA 1729
Qy 651 GTTGATCAGTCTGACCTGTCTAATTAAGAAAGCTTTGATGAGTTGACGAAAGCACTTGAATCC 710
Db 1728 GTTAATAGGATGATTTGACACTTTCAGAGGCTTGAAGAATTTTCGTTTCACTACAGGAG 1669
Qy 711 CTTTCAAGGAGAGAGAGCGATCGGCTAAAGCAAGTGTATAGACATTTTGAATTTCTTTGCAT 770
Db 1668 CTCGAAAAAGAGAGAGCAATAGGTTGGAGAAGGTTCTTGTATATGTAGCATGATACAT 1609
Qy 771 TCCTTATGTGAGGTCCTTGGCATAGATTTCAAGCAAAACAGTATATGAGGTGCAACCTTAGC 830
Db 1608 AATCTCTGTACTGTTTAGGGATGATTTTCTTAGCAGAGTACTGAGTTCATCCAGT 1549
Qy 831 TTGGACGAAAGCTGAAGGA-----TCAAGAACTCTGAGCAACACTACAATTTGAGAGCTT 884
Db 1548 TTGGATGACTCTATTTGGTGACAACCTGCAAGAGCATTAGCAATGATACACTGTCAAACTT 1489
Qy 885 GCTGTCGCGGCAAAACAGACTGGGTGAATTAAGATCCAAAGGATGCAAAAGCTTCAGAT 944
Db 1488 GACAAGACTGTAGCCACACTTAAATGAAGACAAAGAGTCAAGACTAAGCAAGTTCAAGAG 1429
Qy 945 TTTGCTTCTAGCATGCTCGAGCTATGGAATCTCATGGATCTCCTCACTTCAAGAGCAGCAG 1004
Db 1428 CTTGCTGTGCTCAACTCTAGCATCTTTGGGATCTCATGGATGCCCCCATGCAAGAAAGGAGC 1369
Qy 1005 ATGTTTTCAGAAATATAACATGCAATATTTGCTGCTTCAGAAACAGAGATTAATGCAACAAAC 1064
Db 1368 ATGTTTGTGATCACTGCTGCAACAGATCAGCTCTGTAGACAAAGTCACTGCACTGGA 1309
Qy 1065 ACCCTCTCCACAGATTTCTGAAATTTATGTCGAATCTGAGGTGTTAAGGCTTGAAGCACTG 1124
Db 1308 GCTCTTGGCTTGGATCTCAATTGAGCAAGCTGAGGTAGAGTTTCAAAAGGCTAGACAGCTA 1249
Qy 1125 AAAGCAAGTAAGATGAAGATCTGTTTTTAAAGAAAGAGAGAGAGAGTGAAGAGCAGATAGA 1184
Db 1248 AAATACAGCAAGATGAAGAAATAGCTTTCAAGAGCAAACTGAGTTGGAAGACATCTAT 1189
Qy 1185 AGACGTGCTCATCTTTGTCGAGGAAAGGTTATGCAAGAGGAGTTTATGCAATTTGAAGCTATT 1244
Db 1188 GCTGTGCTCATGTAATAGATACAGCTGCTGCTGCCATGAAAAATAATTTGGCGCTGATT 1129
Qy 1245 GAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTCACATGCAACA 1304
Db 1128 GAGGACGAGCAACATAGAACCACTAGAAATTTGATTCAGATATGGAAGAGCCAGATATCAAAA 1069
Qy 1305 GTGAAAGAGAGAGCTTTTAGCCGAGGATATTTCTTGAGAAAGTTTCAAGAGATGCGCAAAAT 1364
Db 1068 GCAAAGGAAGAGAGCTTTGATGAGAAAGAAATATTAGACAAAGTTGAAAGATGAGATGCTCT 1009
Qy 1365 GCTTGTGAAGAGAGAGCTTGGCTTGAAGATTTACAAAGAGATGATTAATCGTTCAATGCT 1424
Db 1008 GCTTGTGAAGAGAGAGCTTGGCTTGAAGATTTAACCAGGATGACAAACAGATATACTCA 949
Qy 1425 GAGAGGGAGCAGCTCTAAACAATAAGAGGCTGAAAAGGCTCGTACTTGTGTCACAAAG 1484
Db 948 AGCCGAGGTGCCACCTGAATCTCAAACTGCTGAAAAGGCTCGGATTTCTTGTCAACAAG 889
Qy 1485 ATTCTCTGGAATGGTAGATCTTTTGAACAACAAATTTGCTGCAATGGAAGAAATGAACGAGGA 1544
Db 888 ATTCAGCACTTGTGTTGAAACTCTGTTGGCAAGACCCAGGCGATGGGAGGAGAGCCGTGG- 830

Qy	1545	AAGGAGATTTCACATATGATGGTTGTAGCCTTTCGTCATGCTTGATGAATATATGTTCC	1604
Db	829	--TCTGCTCTTTATGATATGATGGTGTACCTCTCTTAGCTATGTTGGATGAATATGTCATG	772
Qy	1605	GTTTCGTCAGGAGAAAGCAGAGAGAAAGAGACAAAGGGATCAGAG---AAGCTCCAG	1661
Db	771	CTTAGCGGAGAAAGGGAAGAGATAGAAAGATAGAGGGAACAAAAGCGCTACATCGAG	712
Qy	1662	GATCAGCTCAAAGCGGACGAGGAGCTTTTGTCGGATCAAAAACCCAGTCCATCCAAAGCCC	1721
Db	711	CAACACTGAACACCGACCATGAAGGCCCATTTGGTTCAGCTGTGACCCCTAACAGACCA	652
Qy	1722	CTAAGTACAAGAGGCACCTAG	1744
Db	651	GTCAGTGCAAGAGGTACCTGG	629

RESULT 12

```

US-10-424-599-111490/C
; Sequence 111490, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111490
; LENGTH: 2712
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71686C.1
US-10-424-599-111490

```

Query Match	15.8%; Score 389.6; DB 16; Length 2712;
Best Local Similarity	55.0%; Pred. No. 7.8e-99;
Matches	842; Conservative 0; Mismatches 669; Indels 21; Gaps 3;
QY	238 CGACATGCGGATTTCGCTTCTACTGGAGCTCAATGTGATTTCGGGATGAGGTCCGTGAGCCGC 297
Db	2241 CTACTTGCACTGCTTTGCTCAGAGAACTTGACCAATATGGAACGATATTGGGGAGACC 2182
QY	298 ACACGAGGAGGACAGATGCTGCTGGAGCTCGACGACGAGTGCTGCTGGAGGTTCTACAGGC 357
Db	2181 AAGTGGAGAAAGATTCGAATCTGATGAGCTGGAGAGGGGAATGCTTAGAAGTATACAGGA 2122
QY	358 GGAAGTGCACGAGGCGAACCGAGCGCGCCACAGCTGCGGAAAGGCCATCGCCGAGGGCG 417
Db	2121 GAAAGGTTGATGAGCTGCTTAAACCAACGACGCTTTTCAATCAACGTTGCCGCCCAAGG 2062
QY	418 AGGCAGAGCTCGCCGGCATCTGCTCAGCCATGGCGGAGCCGCCCGTSCA---CGTTAGAC 474
Db	2061 AAGCAGAGCTTGCTACACTGATGGCTGCACTTGGAGAACATGATATTTCATTCCACGATTA 2002
QY	475 AGTCAATCAGAGCTTCATGGCTTAAGAGAGGAGTTGAATGCAATGTGTCCGATTTTGG 534
Db	2001 AGATGGAGAAGAGATCAGCATCCCTGAAGCAGAAACTTGGCGTCCATCACACCTTTGGTTG 1942
QY	535 AAGAAATGAAAAAGAAAAAGGTCCAGCATGGAACAGAGTTTGTTCATGTCATAGAGCAGA 594
Db	1941 AAGAACTGAAGAAGAAAAAAGATGAAAGATTGAAGCAATTTGAAGATGTAAGACATCAA 1882
QY	595 TTAAGAAATTTTCGTCGTAATTAAGGCCGCCGATTTTGTTC-----CCTTTA 642
Db	1881 TAGAAAAGATAAGTGGAGAGATTTTTGGATTCCATTCTGTCAATAATGCTTTGAGCAGCA 1822
QY	643 AAGTTCGGGTGATCAGCTCTGACCTGTCATTAAAGAAAGCTTTGATGAGTTGACCAAGGACC 702

[illegible]

Db 741 AAAATAACAGCTTTTAGAAAGACTAATAGTTAC 710

RESULT 13

US-10-425-115-39061
 ; Sequence 39061, Application US/10425115
 ; Publication No. US20040214272A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53222)B
 ; CURRENT APPLICATION NUMBER: US/10/425,115
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 369326
 ; SEQ ID NO 39061
 ; LENGTH: 2422
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: MRT4577_135624C.1
 US-10-425-115-39061

Query Match 15.4%; Score 379; DB 18; Length 2422;
 Best Local Similarity 55.3%; Pred. No. 7.3e-96;
 Matches 830; Conservative 0; Mismatches 650; Indels 21; Gaps 4;
 QY 237 ACGACATGGGATTCGCTTCTACTGGAGCTCAATGTGATTTGGATGAGTCGGTGAGCCC 296
 Db 552 ACGACATGGGATTCGCTTCTACTGGAGCTCAATGTGATTTGGATGAGTCGGTGAGCCC 611
 QY 297 GACACGACGAGGACAGGATGCTGCTGGAGCTCGACGAGGAGTGCCTGGAGTCTACAGG 356
 Db 612 GATGAAGATCGTGACAAAGATTTACTTTCAGCTAGACGAGAGTGCCTTAGATGTTATAGG 671
 QY 357 CGGAAGGTGACGAGCGGAAACCGGAGCCGCGCCAGCTCGGAGGCCATCCCGAGGCC 416
 Db 672 AGAAAGGTTGATCAGGCTTCTAGCTCCAGGTCTATGCTCTCCACAGCTTTCCCACTCC 731
 QY 417 GAGGACAGCTCCCGGCTATGCTCAGCATGGGCGAGCCGCGCTGACGCTTAGACAG 476
 Db 732 AATCGGAAGTACTGATGATCTCTTCTGATGATGAGGAGTCTATTTCTGCTGCTC--- 789
 QY 477 TCAATCAGAGCTTCATGGCTTAAGAGAGGAGTGAATGCAATGCTTCGGTATTGGAA 536
 Db 790 -CTGACAAGACATCTGGTACNATCAAGGAGCACTAGCAGCTATATCACCATCTTTAGAG 848
 QY 537 GAAATGAAAAGAAAAGGTGCAACGATGGAACCACTGTTGTCATGCTATAGAGCAGATT 596
 Db 849 ACACCTCCGAGGAAAAGAGAGTGGGTGAAAGAGTTTGCTCATGTACAATTCAGATT 908
 QY 597 AGAAAAATTCGCTGCAATTAAGCCAGCCGAGATTGTTCCCTTTAAAGT-----T 647
 Db 909 CAGTACTCTGATGAAATTCGCTGGGATCTACATATTGGTGAAACAGTTGGAAACACCC 968
 QY 648 CCGGTTGATCAGTCTGACCTGTGCTTAAGAAAGCTTGATGAGTTGACGAGGACCTGGAA 707
 Db 969 CATGTTAATGCGGATGATCTTTTCAGTGAGGAGATTAAATGAAATATCTTTCTGAATACAA 1028
 QY 708 TCCCTTCAGAGAGAGAGCGATTCGCTTAAACCAAGTGTAGAACTTTGAAATTTCTTTG 767
 Db 1029 GCCTCGCAAAAGAAAAGAGTACTAGGCTCTCAAGGATTTCTGAATCCGAGTTCAGTG 1088
 QY 768 CATTCTTATGTCAGGCTTGGCATAGATTTTCAAGCAACAGTATATGAGTGCACCTT 827
 Db 1089 CAGATCTCTGTTCTGTTACTTGGCATGAGATTTTGTGGGCACTGTTACTAAAGTTCAATCA 1148
 QY 828 AGCTTGACGAAAGCTGAAG-----ATCAAAGAACCTGAGCAACACTACAATTTGAGAG 891
 Db 1149 AGTCTTGATGATCTGTTGGTGTGCTCAATCCAAAGGATTTAGTGGCGAAACAAATATCCAG 1208

QY 882 CTTGCTGCTGCCCAAAACAGACTCGCTGAAATGAAGATCCAAAGATGCCAAAAGCTTCAA 941
 Db 1209 CTGCTAAAATGGTGTGTTACTTTGAAGAGAAAATCGAAGAGGTTTGCAAAAGATTCAA 1268
 QY 942 GATTTGGTCTAGCATGCTGAGCTATGGATCTCATGGATCTCCACTTTGAAGAGCAG 1001
 Db 1269 GCTCTAGCTTCCAGCTATCTGATCTTTGGAACTTTAATGGATTTCTCTGTGGAGAACAG 1328
 QY 1002 CAGATGTTTTCAGAAATATAACATGCAATATTCTGCTTCAGAAACAGAGATTAACCTGAACCA 1061
 Db 1329 CAACCTTTTCTGCTAGCTAACATGATATGCTCTCAACATTTGGATGATGTGATAGTTCCA 1388
 QY 1062 AACACCTCTCTCCACAGATTTTCTGAAATATGTCGAATCTGAGGTTTGAAGGCTTGAACNA 1121
 Db 1389 GGAGCTCTGGCTCTTGTATGTAATTTCAGCAGCGGAACTCGAAGTTTGAAGGCTTGTATCAG 1448
 QY 1122 CTGAAGCAAGTAAAGTAAAGATCTTTTAAAGAAAGAGCAGAACTAGAAGAGCAT 1181
 Db 1449 CTAAGAGCTAGTAGGATGAAGGATATTGCAATTCAGAAACACAGCTGAACCTTGAAGATATA 1508
 QY 1182 AGAAGACGTGCTCATCTTTTGGCGAGGAGGTTATGCAAGAGGAGTTTAGCATTTGAAGCT 1241
 Db 1509 TATGCCCGGGCTCATGTTGCAAGTAGATTCTAGTGTGCAAGAGATAGGATATGCTCTATT 1568
 QY 1242 ATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTTGAACAAATTTGAAGCTCACATTGCA 1301
 Db 1569 ATTGAGTCCAGTAGTTTCGAGCCTTCAGAACTACTGCTGTGATATGGAGAACCCAGATCTT 1628
 QY 1302 ACAGTGAAGAGGAGCTTTTACCGGAGGAGTATTTCTTGAAAGAGTTTGAAGAGATGGCAA 1361
 Db 1629 AAAGCAATGAGGAGTCCCTTAAGCAAGAAAGGACATATTGGAGAGGGTTGATAGGTGGATG 1688
 QY 1362 AATGCTTGAAGAGGAGGAGCTGCTGGAAGATTAACAACAAAGATGATAATCGTTACAAT 1421
 Db 1689 TCAGCATGTGAAGAAAGAAAGCTGGCTTGAAGACTACGCGGACGATTAACAGGTACAGT 1748
 QY 1422 GCTGGAGGGGAGCAGCATCTAACACTAAAGAGGGCTGAAAAGGCTCGTACTTTTGGTCAAC 1481
 Db 1749 GCAACTCGGGGTGCACATCTGAATCTGAAGCGTGCAGAAAGGCGCTGTTTGGTCAAT 1808
 QY 1482 AAGATTCCTGGAATGGTAGATGTTTGAACAAAAATTTGTCATGGAAGAAATGAACGA 1541
 Db 1809 AAAATTCGGCTATTGTTTGACACACTGCTTGCAAAAGACCCGGGCTAGGAAACAGAGCAT 1868
 QY 1542 CGAAAGGAGGATTTCACATATGATGCTGTAGCTTTTGTCAATGCTTGTGATGAATATATG 1601
 Db 1869 GG---CGTGGCTTCACTACGACGGCTCCCTCTTCTTGGCATGCTGACGAAATACAAA 1925
 QY 1602 TTTGTTCTGTCAGGAGAAAGAGCAAGAGAAAGAGAGCAAGGGATCAGAAAGAGCTCCAG 1661
 Db 1926 ATCTTGAGGAGGAGAGGAG 1985
 QY 1662 GATCAGCTCAAGCGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1721
 Db 1986 GAACAGCTAGCGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2045
 QY 1722 C 1722
 Db 2046 C 2046

RESULT 14
 US-10-425-114-35072
 ; Sequence 35072, Application US/10425114
 ; Publication No. US2004003488A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaka, Jack E.
 ; APPLICANT: Cao, Yongwei

Query Match		14.0%; Score 346.4; DB 17; Length 2345;
Best Local Similarity		54.1%; Pred. No. 1.2e-86;
Matches 826; Conservative		0; Mismatches 676; Indels 24; Gaps 5;
Qy	237	ACGACATGCGATTGCTTCTACTGAGAGCTCAATGTGATTTGGGATGAGTTCGGTAGCCCC 296
Db	373	ACTTCGTGTGGATCATTTGCTGCAAAACTCGAGTCGGATCGGTATGGATGCAAGTTGGTGAAGC 432
Qy	297	GACACGACGAGGACGAGATGCTGTGGAGCTCGAGCAGGAGTGTCTGGAGTCTACAGG 356
Db	433	GAGGAGGACCGGACAAAGGTTCTATTTCAGCTGGACGAGATGTTTGGATGCTACAG 492
Qy	357	CGGAAGGTCCACGACGCAACCGGAGCCGCGCCAGCTCGGGAAGGCCATCCCGAGGGC 416
Db	493	CGGAAGTTGATCAGCAACCAACAAATCAAGGGATCTTCTCCAGGCACTGGATTACTCA 552
Qy	417	GAGGACAGACTCGCCGGCATCTGCTCAGCCATGGGGAGCCGCCGTGCACGTTAGACAG 476
Db	553	AAGACAGAGCTTGCTAGGCTTGCTTCTGCCCTTGGTGAATAATCCATAGATATAAGTC-- 610
Qy	477	TCAATCAGAGCTTCATGGCTTTAGAGAGGAGTTGAATGCAATTCGTCGGTATTGGAA 536
Db	611	-CTGAAACACGACGCGACTATCAAGGAACAACTAACTGCTATAGCTCCAACTTGAA 569
Qy	537	GAAATGAAAAGAAAAGGTGCAACGATGGAACCACTTTTGTTCATGTCATAGAGCAGATT 596
Db	670	CAATTGGGCAGNAGAAAAGAGAGATAAAGGAATTAGCTAATATACAGTCAAGAAATT 729
Qy	597	AAGAAAATTCGTCTGAATAAAGGCCAGCGATTTTGTTCCTTTAAAGTTCC----- 649
Db	730	GAGCAATACGTGGTGAGATTGCTGTAATCTGAGATGGGACAGCAAGTGGCACTPACCA 789
Qy	650	--GGTTGATCAGTCTGACCTGTCTTAAGAAAGCTTGATGAGTGAAGAGCACTGGAA 707
Db	790	CAGATTAAATGAGGATGATTGACAGTTCGGAAGCTTCGAGAGTTCCAGTTACAGCTTCAA 849
Qy	708	TCCTTTCAGAGGAGAGAGCGATCGCTAAAGCAAGTGATAGAACATTTGAATTTCTTTG 767
Db	850	GAGCTTGAGNAGAGAGAGTCTGAGCTGGAAAGGTTCTTGAGCATGTGTGATGGTA 909
Qy	768	CATTCTCTATGTAGAGTGTGGCATAGATTTCAAGCAAAACAGTATATAGGTGACCCCT 827
Db	910	CATGATCTATGCAATGTCTGGATGGATTTTATGAACAATAAATACTCAAGTCCATAGC 969
Qy	828	AGCTTGGACGAGCTGAAGS-----ATCAAGAACTGAGCAACACTCAATTTAGAGG 881
Db	970	AGTCTAGTAGTCTATTGGTAA CGAAACA CAAGCAATAGCAATGAAACCCCTGTCAAAA 1029
Qy	882	CTTGCTGCTGCCGCAACAGACTGCGTGAATGAAGATCCAAAGGATGCAAAAGCTTCAA 941
Db	1030	CTTGATAGGCCAATTGGTACTCTTAATAGGACAAAGGTTGAGACTTGGAAGCTTCAA 1089
Qy	942	GATTTTCTTCTAGCATGTCTGAGCTATGGAATCTCATGGATATCTCCACTTGAAGAGCAG 1001
Db	1090	GAGCTTGCCACTCAACTCTATGATCTATGGATCTTATGGACACCCCGTGGGAAGAAAG 1149
Qy	1002	CAGATGTTTCAGAAATATAACATGCAATATGCTGCTTCAGAA CAAGAGATACTGAACCA 1061
Db	1150	AGCTTGTGTTGATCATGTTCTGTCNAATAGAACAGCAACTGTTGAGGAGGTCTGTTTCT 1209
Qy	1062	AACACCTCTCCACAGATTTCTCGAATATGTGCAATCTGAGGTGTAAAGCTTGAACAA 1121
Db	1210	GGTGCACTTCTGTTGATGTTGATGATCAAGCACAGACTGAGGTTGAAAGTTGGATCAG 1269
Qy	1122	CTGAAAGCAAGTAAAGATCTTTGTTTAAAGAAAGCAGAACTAGAAAGCAT 1181
Db	1270	CTAAATACAGCAAGATGAAGAAATAGCTTTTAAAGAAAGCAAGCACTCTGGAAGATATT 1329
Qy	1182	AGAAGACGCTCTACTTGTGGGAGGAAAGTTATGACAGAGGATTTACATTTAGACT 1241
Db	1330	TATGCTAGCACTCATGTTGTTATTAGACACAGCAGTGTGCCCATGAGAAATACAGGCAC 1389
Qy	1242	ATTGAAGCTGGAGCTATTGATCCCTCACTAGTACTTGAACAAATTTGAAGCTCACATTGCA 1301

Search completed: November 20, 2004, 00:37:52
Job time : 1219 secs

Db	1390	ATTGAGTCAGGGAACATGGAACCTTTCAAGACTGATTTGCTGACATGGATAGTCAGATACTG 1449
Qy	1302	ACAGTGAAGAGGAGGCTTTTAGCCGGAAGGATATTCTTTGAGAAAGTTTGAAGATGGCAA 1361
Db	1450	AAAGCAAGGAGGAGGCTCTGAGCAGAAAGAAATATTAGATAAAGTTGAGAGATGGATA 1509
Qy	1362	AATGCTTGTGAAGAGGAGGCTGCTGGAAGATTACAACAAAGATGATAATCGTTACAAAT 1421
Db	1510	TCCTCATCGAGGAGGAAAGCTGGCTTCAAGACTATAGCAGAGATGACAAACAGGTATAAC 1569
Qy	1422	GCTGGAGGAGGACACATCTAACACTAAGAGGCTGAAAGGCTCGTACTTTTGGTCAAC 1481
Db	1570	TCAGGCCAGGTGCTCACCTGAATCTCAACGCTGAGAAAGGCCGCTATTCTTAGTCAGC 1629
Qy	1482	AAGATTCTGGAATGCTAGATGTTTGAAGAACAAAATTTGCTGCAATGGAATAATGAACGA 1541
Db	1630	AAAGTTCCAGCCCTTGTGAAACTCTGGTGGCTAAGACGAGAGCTTGGGAAGAAAATCAC 1689
Qy	1542	GGAAAGGAGGATTTTCAATATGATGCTGTAGCCCTTTCGTCAATGCTTGTGATGAATATG 1601
Db	1690	GG---GCTGCCATTTATGATGACCGTGTCTCTGCTAGCAATGCTGGATGAGTATGTC 1746
Qy	1602	TTGCTTCTCAGGAGAAAGCAGAGAGAGAGACAAAGGGATCAGAAAGAGCTCCAG 1661
Db	1747	ATTCTTAGGCAAGAGAGGGAAGAGAGAAAGAAATGCGGGAAACAAAACGCCAGACG 1806
Qy	1662	GATCAG---CTCAAAGCGAGCAGGAAGCTTTGTACGGATCAAAACCCAGTCCATCCAAAG 1718
Db	1807	GAAAGCTTACTTAACATCGACCGTGGGGCCATTTGCGNACACGGGTCAATCCCTACAGA 1866
Qy	1719	CCCCTTAAGTACAAAGAGGCCACCTAG 1744
Db	1867	GTGACCAGTGCACAAAGAGGTGGCTGG 1892

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 22:08:20 ; Search time 101 Seconds
(without alignments)
3930.777 Million cell updates/sec

Title: US-10-619-685-2
Perfect score: 3496
Sequence: 1 MSSAVKQQLHQMSITTCDSLL.....DIDYFERRLAIFYLARQWV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 02:*

1: uniprot_prot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1929.5	55.2	730	Q6L4A9	G6L4a9 solanum dem
2	1299.5	55.2	730	AAT40494	Aat40494 solanum d
3	1900.5	54.4	707	Q9FHM4	Q9fhm4 arabidopsis
4	1625.5	46.5	645	Q8LQE3	Q8lqe3 oryza sativ
5	1362.5	39.0	648	Q9L2Y0	Q9l2y0 arabidopsis
6	1314.5	37.6	587	Q9FLP0	Q9flp0 arabidopsis
7	1311	37.5	580	Q9FEV9	Q9fev9 nicotiana t
8	1296	37.1	576	Q84VU1	Q84vul daucus caro
9	1279.5	36.6	584	Q9FEV8	Q9fev8 nicotiana t
10	1268.5	36.3	582	Q9FEV7	Q9fev7 nicotiana t
11	1236.5	35.4	578	Q8LEG3	Q8leg3 arabidopsis
12	1208	34.6	608	Q9SIS3	Q9sis3 arabidopsis
13	1203.5	34.4	603	Q8L836	Q8l836 arabidopsis
14	1202	34.4	511	Q9LVB1	Q9lvb1 arabidopsis
15	1182.5	33.8	589	Q6ZII3	Q6zii3 oryza sativ
16	1182.5	33.8	589	BAD12872	Bad12872 oryza sat
17	1181.5	33.8	559	Q6Z2N0	Q6z2n0 oryza sativ
18	1181.5	33.8	559	BAD08014	Bad08014 oryza sat
19	1168.5	33.4	635	Q9LQW1	Q9lqw1 arabidopsis
20	1127.5	32.3	607	Q8H006	Q8h006 oryza sativ
21	1066	30.5	570	Q6I569	Q6i569 oryza sativ
22	1042	29.8	587	Q9ZVJ3	Q9zvjj arabidopsis
23	1004	28.7	473	Q9SZ16	Q9sz16 arabidopsis
24	966	27.6	626	Q852G8	Q852g8 oryza sativ
25	905	25.9	592	Q9C7G0	Q9c7g0 arabidopsis
26	878.5	25.1	488	Q6ZJ66	Q6zj66 oryza sativ
27	878.5	25.1	488	BAD09028	Bad09028 oryza sat
28	354	10.1	609	Q7T0Z3	Q7t0z3 xenopus lae
29	334	9.6	605	Q802Y9	Q802y9 brachydanio
30	334	9.6	605	Q6DI26	Q6di26 brachydanio
31	331	9.5	603	Q99K43	Q99k43 mus musculus

32 323.5 9.3 620 2 O43663 O43663 homo sapien
33 307 8.8 516 2 Q8CE25 Q8ce25 mus musculu
34 294.5 8.4 581 2 Q7ZXD0 Q7zxd0 xenopus lae
35 280.5 8.0 576 2 Q9BSB6 Q9bsb6 homo sapien
36 275.5 7.9 598 2 Q6DI02 Q6di02 brachydanio
37 275.5 7.9 845 2 Q8FKN0 Q8fkn0 candida gla
38 271.5 7.8 801 2 Q6CJS4 Q6cjs4 kluyveromyc
39 267.5 7.7 407 2 Q6L5I2 Q6l5i2 oryza sativ
40 267.5 7.7 407 2 AAT39217 Aat39217 oryza sat
41 264.5 7.6 595 2 Q7ZVK0 Q7zvk0 brachydanio
42 262 7.5 579 2 Q6DI23 Q6di23 brachydanio
43 259.5 7.4 885 1 ASE1 YEAST P50275 saccharomyc
44 254.5 7.3 888 2 Q7SEEA Q7seea ashbya gosb
45 254.5 7.3 888 2 AAS50503 Aas50503 ashbya go

ALIGNMENTS

RESULT 1

Q6L4A9
ID Q6L4A9 PRELIMINARY; PRT; 730 AA.
AC Q6L4A9;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Putative microtubule-associated protein.
GN Name=PGEC561.15;
OS Solanum demissum (wild potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=50514;
RN [1]_TaxID=50514;
RP SEQUENCE FROM N.A.
RA Buell R.C., Liu J., Zaborsky J., Tallon L., Baker B.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Ronning C.M.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
DR EMBL: AC136471; AAT40494.1; -;
DR InterPro: IPR007145; MAP65_ASE1.
DR Pfam: PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 730 AA; 82876 MW; E4E80EBD8F60D4E5 CRC64;

Query Match 55.2%; Score 1929.5; DB 2; Length 730;
Best Local Similarity 57.1%; Pred. No. 1.5e-79;
Matches 418; Conservative 91; Mismatches 154; Indels 69; Gaps 15;

QY 12 MSTTCDLLELNLVNWDEVEGPDTRDRMLLEQECLEVYRKVDQANRSALRAKAIA 71
DB 1 METTCGLLHLELQIINWDEVEGCDTERDRMLPEIERECLEVYRKVDQANRSALRAKAIA 60
QY 72 EGAEALAGICSAEGEPVIVRQSKLHLGRLELNAIVPYLEEMKKKVKVWVHVIE 131
DB 61 DSEAEALAIICSAANGRPVHIKQSDKSQGLKAEALRAVLPELSEMRKRDKNQFIEVMK 120
QY 132 QIKTISSEIRPADFPVPKVPVQSDLSLRKLDLTLDLESLOKESKDRKLQVTEHLSLH 191
DB 121 QITKINEI--YRFTSASLVDESLSLRKLELHTELHTLQEKSERLKQVNLHGLTN 178
QY 192 SLCEVLGIDFKQTVVEVHPSLDEAGSKNLSNTTIERLAAANRLREMIQRMKQLDFA 251
DB 179 SLCSVLGMDFKHTINEVDNLGSEAEKNCDDTIQNLAAATIQRLQEVKLQRMQLDLT 238
QY 252 SSMLELWNLMDTPLEEQQMFQNTCNIAASEQIETPNTLSTDPNLYVSEVLRRLQKA 311
DB 239 TSMLELWNLMDTPLEEQQMFQNTCNIAASEQIETPNTLSTDPNLYVSEVLRRLQKA 298

SEQUENCE 707 AA; 80301 MW; 30EAF1E5BCC6B9A CRC64;
Query Match 54.4%; Score 1900.5; DB 2; Length 707;
Best Local Similarity 55.1%; Pred. No. 3e-76;
Matches 397; Conservative 120; Mismatches 148; Indels 55; Gaps 12;
QY 1 MSSAVKQDLQHQMTTCDLSLLELNVWDEVGSPDTRDRMLLEBOECLEEVYRRKVDQAN 60
DB 1 MASVQKDPLOVETTCGSLFELQIWDVGEETEDRQMLLEBOECLEEVYRRKVDQAN 60
QY 61 RSRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKHLGLREELNAIVPYLEEMKKVKV 120
DB 61 RCRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKHLGLREELNAIVPYLEEMKKVKV 120
QY 121 ERNNOFVHVIEQIKISSIR-PADFPVPEKVPVQSDLSLRKLDLTKLESLOKESDR 179
DB 121 ERNNOFVHVIEQIKISSIR-PADFPVPEKVPVQSDLSLRKLDLTKLESLOKESDR 180
QY 180 LKQVIEHLNLSHSLCEVLGIDFKQTVYVHPSPDEAGSKNLSNTTIERLAAANRLRM 239
DB 181 VETIRKHLCTLSHCSVLGMDNEVGVNPTLSDPEGRSLSDHTIEKLGAAVQKMEV 240
QY 240 KIORMKQLQDFASSMELNLMMDTPLEEQOMFONITCNIAAEOETEPNTLSTDFLNVY 299
DB 241 KIORMKQLQDFASSMELNLMMDTPLEEQOMFONITCNIAAEOETEPNTLSTDFLNVY 300
QY 300 ESEVLRLQKASMKDLVKKAELEHRRRAHLVGEYGAEPFIEAIEAGADPSLV 359
DB 301 EAEVRLDEYKASKMELVKKSELEECRKTHLLPVSDSAIDQTVIAESGIVDATWV 360
QY 360 LEQIEAHIAITVKEEAFSRKDIIEKVERWONACBEEAWLEDYNDKDNRYNAGRGHLLTKR 419
DB 361 LEHLEQHSIKIEEALSRKEILRVEKWLACDEESWLEYNRDDNRYNAGRGHLLTKR 420
QY 420 AEKARTLVNKPQWVDVLRKIAAANKNKGEDFTYDGVSLSSMLDEYMFVROEKEKK 479
DB 421 AEKARTLVNKPQWVDVLRKIAAANKNKGEDFTYDGVSLSSMLDEYMFVROEKEKK 479
QY 480 RORDQKQLQDLKABQEAALYGSFSPSKPLSTKAPRHSNGGA-NRRLSLGGATMQPKT 538
DB 480 RORDQKQLQDLKABQEAALYGSFSPSKPLSTKAPRHSNGGA-NRRLSLGGATMQPKT 539
QY 539 DILHKSVAARKTEETIGLSPSSRGLDLAGPIKKLSFNASTLRETETPRKPPA- --- 594
DB 540 N-----KKADRHNDGALS-NRRGLDLAGLPSRKQSNPSEMQLSPVRPFSPIST 591
QY 595 -----QITPGNSVS-----STPVR---PITNNTDE-----NRTPKTFTA 627
DB 592 TVVASKANIATTTTQQLPKONAVNEISSPATPIKNNILRNLEEKMTMMQTPKNVA 651
QY 628 L--NPKTPMTVTPMQWMTPLSKVATPVSILVYDKPVTLOEIDISFEERRLAIVL 685
DB 652 MIPISTPATVSPMHTAPTFTNN-----ARLMSEKPEV-----VEYSFEERRLAIVL 700

RESULT 4

Q8LQES PRELIMINARY; PRT; 645 AA.
ID Q8LQES
AC Q8LQES
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cytokinesis regulating protein-like.
GN Name=P0478H03.25;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN (1)
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,

RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayaashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003452; BAB92709.1; -;
DR Gramene; O8LQES; -;
DR InterPro; IPR007145; MAP65_ASE1.
DR InterPro; IPR009061; Putativ_DNA_bind.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 645 AA; 70887 MW; 2DE3181F1B8F8ED8 CRC64;
Query Match 46.5%; Score 1625.5; DB 2; Length 645;
Best Local Similarity 61.1%; Pred. No. 7.9e-66;
Matches 361; Conservative 16; Mismatches 61; Indels 153; Gaps 9;
QY 1 MSSAVKQDLQHQMTTCDLSLLELNVWDEVGSPDTRDRMLLEBOECLEEVYRRKVDQAN 60
DB 1 MSSAVKQDLQHQMTTCDLSLLELNVWDEVGSPDTRDRMLLEBOECLEEVYRRKVDQAN 60
QY 61 RSRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKHLGLREELNAIVPYLEEMKKVKV 120
DB 61 RSRAQLRKAIAEAGEAELAGICSAMGEPVHVQSQKHLGLREELNAIVPYLEEMKKVKV 120
QY 121 ERNNOFVHVIEQIKISSIRPADFPVPEKVPVQSDLSLRKLDLTKLESLOKESDR 180
DB 121 ERNNOFVHVIEQIKISSIRPADFPVPEKVPVQSDLSLRKLDLTKLESLOKESDR 180
QY 181 KOVIEHLNLSHSLCEVLGIDFKQTVYVHPSPDEAGSKNLSNTTIERLAAANRLRMK 240
DB 181 KOVIEHLNLSHSLCEVLGIDFKQTVYVHPSPDEAGSKNLSNTTIERLAAANRLRMK 240
QY 241 IORMKQLQDFASSMELNLMMDTPLEEQOMFONITCNIAAEOETEPNTLSTDFLNVY 300
DB 241 IORMKQLQDFASSMELNLMMDTPLEEQOMFONITCNIAAEOETEPNTLSTDFLNVY 300
QY 301 SEVLRLQKASMKDLVKKAELEHRRRAHLVGEYGAEPFIEAIEAGADPSLV 360
DB 299 -----
QY 361 EQIEAHIAITVKEEAFSRKDIIEKVERWONACBEEAWLEDYNDKDNRYNAGRGHLLTKR 420
DB 299 -----
QY 421 EKARTLVNKPQWVDVLRKIAAANKNKGEDFTYDGVSLSSMLDEYMFVROEKEKK 480
DB 317 EKARTLVNKPQWVDVLRKIAAANKNKGEDFTYDGVSLSSMLDEYMFVROEKEKK 480
QY 481 Q-----RDKKQLQDLKABQEAALYGSFSPSKP-----LSTKKAPRHSNGGA- ---A 522
DB 358 HATPRAEGAVNVTDRPATTTGVPAGQPHDARTLARTHNAQAARAGRGACVWDSP 417
QY 523 NRRLSLGGATMQ--PPKTDILHKSVAARKTEETIGLSPSSRGLDLAGI 571
DB 418 NTR---DGRTNQVPP-----RAVRPVEPIESIDSLPSVAGELYLVGM 458
RESULT 5
Q8LZY0
ID Q8LZY0 PRELIMINARY; PRT; 648 AA.
AC Q8LZY0
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)


```
Db 367 ISKAEEAFSRKIDILDRVKKMSACBESWLEDYNRDQNRYSASRGAGHNLKRAEKARIL 426
Qy 427 VNKIPGMDVLRRTKIAAKNKGKEDFTYDGVLSLMDLDEYMFVROEKEKKRORDOKK 486
Db 427 VSKIPAMVDLTVAKTAWBEEHSM-SFAYDGVPLLAMLDYGLRQREBEEKKRLREQKK 485
Qy 487 LDQLKAEQEAALYSGKSPSKPLSTKK--APRHSMGGA---NRRLSL 528
Db 486 VQEQPHVEQESAFSTSPSPARPVSAKTVGPRANNGANGTHNRRLSL 533

RESULT 7
Q9FEV9 PRELIMINARY; PRT; 580 AA.
AC Q9FEV9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-JUN-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1a.
GN Name=map65-1a;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Suspension culture;
RA Smertenko A., Saleh N., Igarashi H., Mori H., Hauser-Hahn I.,
RA Jiang C.J., Sonobe S., Lloyd C.W., Hussey P.J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ289862; CAC17794.1; -.
DR InterPro: IPR007145; MAP65_ASE1.
DR Pfam: PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 580 AA; 65259 MW; 7519865D362F508A CRC64;

Query Match 37.5%; Score 1311; DB 2; Length 580;
Best Local Similarity 48.5%; Pred. No. 1-2e-51;
Matches 280; Conservative 103; Mismatches 158; Indels 36; Gaps 9;

Qy 14 TTCDSLLLELVNIVDEVEGPDTRDRMLLELECELEVVYRKVDQNRSAQRLKRAIAEG 73
Db 14 TTCGSLLOLQIWDVEGETDDEDRNMLLQIDQECLEVDYKRVQDAVKSRHLLQALADA 73

Qy 74 EALAGICSAMGEPVHVVRQSNQKLHGLREELNATVPYLEEMKKKKVERWNOVFVHIEQI 133
Db 74 KVELSLLSALGS-KTYVGIPKTSGTIKESQALAIAPALEKLEWQKDDRIKEFFDVSQSI 132

Qy 134 KTSSEIRPADFPFKVPVQSDLSLRKLDLTKDLESQKESDRLKQVIEHLNLSLHSL 193
Db 133 QKISSEIAGTREQVESLTVDSELSLKLDEFQAQQLQELQKESERLQKVLVSVTHDL 192

Qy 194 CEVLGIDFKQTVYVHPSLDEAG--SKNLSNTTIERLAAAANRLREMKITQRMOKLQDFA 251
Db 193 CAVLGMDFFSTVEVHPSLNDSTGVQSKSISNDTSLSLAKTVLVLKEDKQRLHKLQELA 252

Qy 252 SSMLLELVNMDTPLEEQMFQNTICNIAASEQIETPTNTSLTDFLNVYSEVLRLBOLKA 311
Db 253 TQLIDLVNMDTPEERESLDFHVTNCNISASVDEVAIPGALALDLIEQAEVEVERLDOLKA 312

Qy 312 SKMKDLVLKKALEEHRRRAHLVGBEGYAEBSIEAIEAGADIPSLVLQIEAHATATVK 371
Db 313 SKMKEIAFKQAELEDIYARAHVEIDTEAAREKIMALIDSGNVDPPELLADMNQIVNAK 372

Qy 372 EAFSPRKDILEKVERWQACBEEAWLEDYNDKDNRYNAGRAHLLTKRAEKARTLVNKNIP 431
Db 373 EEAHSRKEILDVKEKMAACBESWLEDYNRDNRNRYNAGRAHLLTKRAEKARILVKNIP 432

Qy 432 GMVDVLRTKIAAKNKGKEDFTYDGVLSLMDLDEYMFVROEKEKKRORDOKKLDQL 491
Db 433 ALVDLSVAKTAWBEEQER-DTTFYDGVPLLAMLDYGLRQREBEEKKRLDQKKFHEQI 491

Qy 492 KAEQEAALYSGKSPSKPLSTKK--APRHSMGGA---NRRLSLGGATMQPKPTDILHKS 545
```

```
Db 492 -SKEETVFGSTSPARPLGPKVTGPR-ANGSANGPTSRRLSLNS-----HONG 538
Qy 546 VRAAKKTEIGTLPSSSRGLDIAGLPIKKLSFNAST 582
Db 539 SRSTWOGKRD-----RPIAPLNYVAMT 562

RESULT 8
Q84VU1 PRELIMINARY; PRT; 576 AA.
AC Q84VU1;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE 65kD microtubule associated protein.
GN Name=map65-1;
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Campanulids; Apiales; Apiaceae; Apioideae; Scandiceae; Daucinae;
OC Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC PubMed=12527379;
RA Chan J., Mao G., Smertenko A., Hussey P.J., Naidrett M., Bottrill A.,
RA Lloyd C.W.;
RT Identification of a MAP65 isoform involved in directional expansion
RT of plant cells.;
RL FEBS Lett. 534:161-163 (2003).
DR EMBL: AJ520103; CAD58680.1; -.
DR InterPro: IPR007145; MAP65_ASE1.
DR Pfam: PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 576 AA; 64771 MW; 1E90E3B570CC6128 CRC64;

Query Match 37.1%; Score 1296; DB 2; Length 576;
Best Local Similarity 47.5%; Pred. No. 5.8e-51;
Matches 275; Conservative 102; Mismatches 156; Indels 46; Gaps 10;

Qy 14 TTCDSLLLELVNIVDEVEGPDTRDRMLLELECELEVVYRKVDQNRSAQRLKRAIAEG 73
Db 14 TTCGSLLOLQIWDVEGESDAERDKMLLQLEQECLEVDYKRVQDAVKSRHLLQSLADG 73

Qy 74 EALAGICSAMGEPVHVVRQSNQKLHGLREELNATVPYLEEMKKKKVERWNOVFVHIEQ- 132
Db 74 QVELHSLMSALGS-KTYVGIPNTSGTIKEQLSALAPTLEQLWKQKDERIKESFSDVSQSI 132

Qy 133 QNDMCQITGTIEKVGSPV-----VDSDSLKLKLBFOAKLQELQKESDRLHK 181
Db 183 VIEHLNLSLHSLCEVLGIDFKQTVYVHPSLDEAG--SKNLSNTTIERLAAAANRLREMK 240

Qy 182 VLEFVSTVHDLCAVLRMDFFSTVEVHPSLDSDSIGVQSKSISNDTSLSLAKTVLALBEEK 241
Db 241 IORMOKLQDFASSMELNMDTPLEEQMFQNTICNIAASEQIETPTNTSLTDFLNVYSE 300

Qy 242 EQLHLQLBELAQKLDLNLMDTSEERSLDFHVTNCNVSASVDEVITPGALALDLIEQAE 301
Db 301 SEVLRLEQLKASKWKDLVLKKALEEHRRRAHLVGBEGYAEBSIEAIEAGADIPSLVL 360

Qy 302 VEVERLDQLKASRMKEIAFKQAELEEEIYVHAHVEIDSEAAERIMALIDSGDVEPAELL 361
Db 361 EQIEAHIAIVKEASRKDILEKVERWQACBEEAWLEDYNDKDNRYNAGRAHLLTKRA 420

Qy 362 DMDNQIVVAKKEALSRLDKVEKMSACEESWLEDYNRDNRNRYNAGRAHLLTKRA 421
Db 421 EKARTLVNKNIPGMVDVLRRTKIAAKNKGKEDFTYDGVLSLMDLDEYMFVROEKEKKR 480

Qy 422 EKARILVKNIPALVLSLVAKTAWBEERG-VSFYDGVPLLAMLDYGLRQREBEEKKR 480
Db 481 ORDQKKLQDLK-ABQEAALYSGKSPSKPLSTKKAPRHSMGGA-----NRRLSLGGATMQP 535
```

Db 481 MKDOKKFHEQTNAEQETAFTARPSPARVGIKKAVGSRANGANGTTPNRRLSINA----- 535
QY 536 PKTDILH---SKSVRAAKTEIEIGTLP-SSRGLDIAG 570
Db 536 -----HONGRSRSINKDGRRESMRAAAPLKEDAGSCISG 568

RESULT 9
Q9FEV8 PRELIMINARY; PRT; 584 AA.
ID Q9FEV8
AC Q9FEV8
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1b.
GN Names=map65-1b;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Suspension culture;
RA Smertenko A., Saleh N., Igarashi H., Mori H., Hauser-Hahn I.,
RA Jiang C.J., Sonobe S., Lloyd C.W., Hussey P.J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ289863; CAC17795.1; -;
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 584 AA; 65886 MW; 9507821B1A3D6835 CRC64;

Query Match 36.6%; Score 1279.5; DB 2; Length 584;
Best Local Similarity 48.9%; Pred. No. 3.3e-50;
Matches 267; Conservative 101; Mismatches 161; Indels 17; Gaps 7;

QY 14 TTCDLSLLELNVWDEVGEPDTRDRMLLEBQCELVYRRKVDQANRSRAQLRAIAEG 73
Db 16 TTGCSLLQQLQIWDVEVGQTDDEWHMLLIQIDRECLDVYKRVQDVAKSRAHLLEALADA 75

QY 74 EAEIAGICSMGEPVHVRSQKHLGRLBELNAIVPLEMKKKVKNWQFVHVEIQI 133
Db 76 KIELCRLLAALGE-KTYAGIPEKASGSIKEQLAALAPALENLWKQEDRVKFEFFVNOAQI 134

QY 134 KXISSEIRPADVPFVKVPVQSDLSRLKDELTKDLESLOKESDRKQVIEHLNSHSL 193
Db 135 EKISSEIAGISEHVNPVKVDESLSLKLDEFOAQLOELQEKSERLHKVLEFVSTIIDL 194

QY 194 CEVLGIDFKQTVYVHPSLDEAEG--SKNLSNTTIERLAAAANRLREMKIORMQKLODPA 251
Db 195 CAVLGLDPFSTITEVHPSLNDSTGVQSOSISNDTSLNSRTVLALMEDKQRLQKLOELA 254

QY 252 SKMLLWNLMDTPLEEQMFQNTCNIAASQEQTEPTNTLSTDFLNVYSEVLRLEQLKA 311
Db 255 TQLIDLWNLMDTPEERSLFDHVCNNSASVDEVSIPGALALDLTQEAQVEVERLDQLKA 314

QY 312 SKMKDLVKKKAELEHRRRAHLVGEEGYAEFEFSIEAIEAGAIQPSLVLEQIEAHIAIVK 371
Db 315 SRMKIAPKQAVLEIEFAHAEIDSEAAQKIMLTIESGNIETPAELLTMDNQIVKAN 374

QY 372 EAFSRKDIKVERWQNAACEEAWLEDYNDKDDNRYNAGRAHLTLKAEKARTLVNKP 431
Db 375 EEAHSRKEILEKVEKMAACEEESWLEDYNDQNRYNASRAHNLKRAEKARTLVNKP 434

QY 432 GMVDVLRITKIAAKWNERGKEDFTYDGVSLSSMLDEVMFVROEKEQKRDOKKLODOL 491
Db 435 ALVDSLIAKTRAEQDR-DTFTYDGVPLLAMLDYMWLRHDEERRLRDQKFFHEQI 493

QY 492 KAEQALYGSKPSKPLSTKK--APR---HSMGGANRLSLGG-----ATMOPPKTDI 540
Db 494 SKEPTTFGSTPSPARPLGTQKVCPRANGSANGNRRSLNSHONGSRSTTKDGRD- 552

QY 541 LHSKSV 546
Db 553 -HSRKV 557

RESULT 10
Q9FEV7 PRELIMINARY; PRT; 582 AA.
ID Q9FEV7
AC Q9FEV7
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1c.
GN Names=map65-1c;
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC lamids; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Suspension culture;
RA Smertenko A., Saleh N., Igarashi H., Mori H., Hauser-Hahn I.,
RA Jiang C.J., Sonobe S., Lloyd C.W., Hussey P.J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ289864; CAC17796.1; -;
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 582 AA; 65474 MW; 28C05FDEAAB9A975 CRC64;

Query Match 36.3%; Score 1268.5; DB 2; Length 582;
Best Local Similarity 48.7%; Pred. No. 1e-49;
Matches 266; Conservative 100; Mismatches 163; Indels 17; Gaps 7;

QY 14 TTCDLSLLELNVWDEVGEPDTRDRMLLEBQCELVYRRKVDQANRSRAQLRAIAEG 73
Db 14 TTGCSLLQQLQIWDVEVGQTDDEWHMLLIQIDRECLDVYKRVQDVAKSRAHLLEALADA 73

QY 74 EAEIAGICSMGEPVHVRSQKHLGRLBELNAIVPLEMKKKVKNWQFVHVEIQI 133
Db 74 KIELCRLLAALGE-KTYAGIPEKASGSIKEQLAALAPALENLWKQEDRVKFEFFVNOAQI 132

QY 134 KXISSEIRPADVPFVKVPVQSDLSRLKDELTKDLESLOKESDRKQVIEHLNSHSL 193
Db 133 EKISSEIAGISEHVNPVKVDESLSLKLDEFOAQLOELQEKSERLHKVLEFVSTIIDL 192

QY 194 CEVLGIDFKQTVYVHPSLDEAEG--SKNLSNTTIERLAAAANRLREMKIORMQKLODPA 251
Db 193 CAVLGLDPFSTITEVHPSLNDSTGVQSOSISDDTSLNSRTVLALTEDKQRLQKLOELT 252

QY 252 SKMLLWNLMDTPLEEQMFQNTCNIAASQEQTEPTNTLSTDFLNVYSEVLRLEQLKA 311
Db 253 TQLIDLWNLMDTPEERSLFDHVCNNSASVDEVSIPGALALDLTQEAQVEVERLDQLKA 312

QY 312 SKMKDLVKKKAELEHRRRAHLVGEEGYAEFEFSIEAIEAGAIQPSLVLEQIEAHIAIVK 371
Db 313 SRMKIAPKQAVLEIEFAHAEIDSEAAQKIMLTIESGNIETPAELLTMDNQIVKAN 372

QY 372 EAFSRKDIKVERWQNAACEEAWLEDYNDKDDNRYNAGRAHLTLKAEKARTLVNKP 431
Db 373 EEAHSRKEILEKVEKMAACEEESWLEDYNDQNRYNASRAHNLKRAEKARTLVNKP 432

QY 432 GMVDVLRITKIAAKWNERGKEDFTYDGVSLSSMLDEVMFVROEKEQKRDOKKLODOL 491
Db 433 ALVDSLIAKTRAEQDR-DTFTYDGVPLLAMLDYMWLRHDEERRLRDQKFFHEQI 491

QY 492 KAEQALYGSKPSKPLSTKK--APR---HSMGGANRLSLGG-----ATMOPPKTDI 540
Db 492 SKEPTTFGSTPSPARPLGTQKVCPRANGSANGNRRSLNSHONGSRSTTKDGRD- 550

541 LHSKSV 546
551 -HSRKV 555


```

547 RAAKTEIEIGTL-SPSS--SRGLDIAGLPIKK 575
539 GSKSRDLSLNKIAPSNIIVANTKODAAFPVSR 570

RESULT 11
Q8LEG3 PRELIMINARY; PRT; 578 AA.
AC Q8LEG3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Microtubule-associated protein MAP65-1a.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurooids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=22088475; PubMed=12093737;
RA Haas B.J., Volfovsky N., Town C.D., Troughan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
[2]
RN SEQUENCE FROM N.A.
RA Brover V., Troughan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085430; AAM62657.1; -.
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 578 AA; 56204 MW; A8544528EPB6C2DE CRC64;

Query Match 35.4%; Score 1236.5; DB 2; Length 578;
Best Local Similarity 44.9%; Pred. No. 2.9e-48;
Matches 257; Conservative 122; Mismatches 166; Indels 27; Gaps 8;

QY 15 TCDSLLLELVINWDEGEPTTDRMLLEQSCLEVYRRKVDONRSRAQLKKAEGE 74
DB 15 TCGLTLQQLQEIWDEGESEDERKLLQLLEQSCLELVYRRKVDONRSRAQLKKAEGE 74

QY 75 AELAGICSAMGEPVHVROSQKHLGLREBELNAIVPYLEEMKKKKVQVNMVQFVH 134
DB 75 VELSNLTALGKES-YIDIPDKTSGTIKEQLSAIPALEQLWQKEERVAFSDVQSQIQ 133

QY 135 KISSEIR-----PADFPVKPVVDQSDLSLRKDELTKDLESQKESDRLKQVLEHLN 188
DB 134 KICEEITAGGLNNGPH-----VVDETSLSKRLDDFORQLQKESDRLKQVLEFVS 186

QY 189 SLHSLCEVLGIDFKQTVYVHPSLDEAG--SKNLSNTTIERLAAANRLREMKIORMOK 246
DB 187 TVHDLCAVRLDFLSTVTEVHPSLDEANGVQTKSINETIARLAKTVTLTKEDMQRLLK 246

QY 247 LQDPASSMLLEMLNMDTPLEEQMFQNTICNIAASEQEIPTNTLSTDFLNYVESEVRL 306
DB 247 LQELATQTLTDLNMDTSDSERELFQHVTSNISASVHEVTASGALDLLEQAEVEVDRL 306

QY 307 EQLKASRMKDLVLKKAELBEHRRRAHLVGEEGYAEFEFSTEAGDAIDPSLVLEQIEAH 366
DB 307 DQLKSSRMKEIAFKQSELEIYARAHIEIKPEVVRERIMSLIDAGNTEPTTELLAMDQS 366

QY 367 IATVKEAFSRKQILEKVEWQACEEAWLEDYNDKDNRYNAGRGAHLTKAEAKRTL 426
DB 367 IAKAKEAFSRKEILDVKEWQACEEAWLEDYNDKDNRYNAGRGAHLTKAEAKRTL 426

QY 427 VNKPIGHVDVLRKIAAQNKRGEKEDFTYDGVSLSSMLDYMFMVFRQEKQRORDQKK 486
DB 427 VSKITANVDTLIAKTRAEWSENSM-SPEYDGVPLAMLDYETWMLRQEREKRLKEQKK 485

QY 487 LQDQLKAEQALYGSKPSKPLSTKCAPRHSMGGANRRSLGGATWQPKTDLILKSV 546
DB 486 QOEQPHTDQESAFGSKPSARPVSAKKP-----VGTRVNGGLNETPMRRLSMNSQN 538

547 RAAKTEIEIGTL-SPSS--SRGLDIAGLPIKK 575
539 GSKSRDLSLNKIAPSNIIVANTKODAAFPVSR 570

RESULT 12
Q9SIS3 PRELIMINARY; PRT; 608 AA.
AC Q9SIS3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein At2g01910.
GN Name=At2g01910;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurooids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007069; AAD21782.1; -.
DR PIR; F84430; F84430.
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
KW Hypothetical protein.
SQ SEQUENCE 608 AA; 69406 MW; 5AE1314D0D311413 CRC64;

Query Match 34.6%; Score 1208; DB 2; Length 608;
Best Local Similarity 39.9%; Pred. No. 6e-47;
Matches 256; Conservative 131; Mismatches 185; Indels 70; Gaps 9;

QY 13 STTCSLLLELVINWDEGEPTTDRMLLEQSCLEVYRRKVDONRSRAQLKKAIE 72
DB 15 NTTNNLLRELQKIWEIGEETETEDKRLMLELERECLEQIYQKRVDAANSAKHLQSVAS 74

QY 73 GAELAGICSAMG---EPVHVROSQKHLGLREBELNAIVPYLEEMKKKKVQVNMVQFVH 128
DB 75 IEAEVASLMAALGVNLNINSPKLDKGSK---SLKEKLAATVPLVELRLQKEERMKQFSD 131

QY 129 VIEQIKKISSEIRPAD-----FVPPKVPVVDQSDLSLRKDELTKDLESQKESDRLKQV 183
DB 132 IRAQIEKISGEISGYSDHLNKKAMNISLTLEEQDLTLRLNLEYQTHLRTLQKESDRLNKY 191

QY 184 IEHLNLSLCEVLGIDFKQTVYVHPSL--DEAGSKNLSNTTIERLAAANRLREMKI 241
DB 192 LGYVNEHALCGVLGVDFSQTVSAVHPSLHRTDQEQSTNISDSTLEGLEHMIQKLERK 251

QY 242 QRMVKLQDFASSMLELWNLMDTPLEEQMFQNTICNIAASEQEIPTNTLSTDFLNYVES 301
DB 252 SRQKLVKQVVASLFEUWNLMDTPQEDRTFKGVTVYVVRSEANITFPGILSTETIEQVST 311

QY 302 EYVLRLEQLKASRMKDLVLKKAELBEHRRRAHLVGEEGYAEFEFSTEAGDAIDPSLVLE 361
DB 312 EVDLSLKLKASRMKELVMKRRSELEDLCRLTHIQPDTSTSAEKSTALIDSLGVDPSSELLA 371

QY 362 QIEAHATVKEAFSRKQILEKVEWQACEEAWLEDYNDKDNRYNAGRGAHLTKAE 421
DB 372 NIEMQINKIKDEAQSKQIMDRWLSACEENWLEYNLDENRYNSAGRGHVNLRKAE 431

QY 422 KARTLVNKPIGHVDVLRKIAAQNKRGEKEDFTYDGVSLSSMLDYMFMVFRQEKQKQ 481
DB 432 RARVTINKIPGMVDTLTKTLVWEEDMQK-SFLYDGVRLVNLIEDYKLTFRKQEEKKRY 490
```

```
QY 482 RDQKLODQKAOEALYSGSPSKPLSTKTKAPRHSMGANRRLSLGGATWQPPKTDIL 541
Db 491 RQKQKQDQLLQFRESYSGSPSKPRSSFRKPNFGFNISNGN----- 532
QY 542 HSKSVRAAKTBEIGTLPSSSRGLDIAGLPIKLSFNASTLRETTETPRK-----P 592
Db 533 -----GSVPPTPRRG-----SVGTTTPDVLTPRSYSGHRQNGY 567
QY 593 FQAITPGNSVSTPVRPTNTNNTEDENRTPKTFPTALNPTPM 634
Db 568 FKFEV---RRLTPTPLNYVAMQKEDTVSTTYTSYSSEPD SPL 606

RESULT 13
Q8L836 PRELIMINARY; PRT; 603 AA.
AC Q8L836;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein Atg14690.
GN Name:Atg14690;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Tripp M., Miranda M.,
RA Palm C.J., Bowser L.K., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayaishizaki Y., Iehida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RA Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Dale J.M., Hayaishizaki Y., Hsuan V.W., Iehida J., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY120768; RAME3326.1; -.
DR EMBL; BF008373; AAP3732.1; -.
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
KW Hypothetical protein.
SQ SEQUENCE 603 AA; 69060 MW; 7C2B37A23CBB889D CRC64;

Query Match 34.4%; Score 1203.5; DB 2; Length 603;
Best Local Similarity 46.2%; Pred. No. 9.6e-47;
Matches 246; Conservative 112; Mismatches 148; Indels 27; Gaps 6;

QY 13 STTCDLLELVINWDEYGEPTTDRMLLEQCELVYRRKVDQANRSRAQLKRAE 72
Db 15 NTTNALLRELOKIWDIGESDAEKRMLEKECELIYRRKVDQANRSRAQLHQSLVS 74
QY 73 GEAEALAGCSAMG-----BPPVHVRSQNKHLGRLAEALNAPYLEEMKKKVRWQFVH 128
Db 75 IEAEATSLAALGVFNESHSPKAKESGK---SLKEKLAARVPMLEDLRLQDKRMKQFVD 131
QY 129 VIEQIKKTSSEIRPADFPFKVPV-----DQSDLSLRKLDLTVDLESLOKESDRLKQV 183
Db 132 IKAQIEKNSGISYSGDQNKNTWVGSALDQDULTRKLNYSQTHLSRLOKESDRLNKV 191
QY 184 IEHLNLSHLCVGLIDFKQTVYVHPSL---DEAGSKNLSNTTIERLAAANRLREMKI 241
Db 192 LDYNEVHTLCVGLVDVGQIVSEVHPSLHRTDHEQSTNISDDILDGLHMHHLKTKERS 251
QY 242 QRMKQLODFASSMLELWNLMDTPLEEQQMFQNTCNIAASQEITEPNTLSTDFLNYVES 301
```

```
Db 252 VRFQKLDVAGSLFELWNLMDTSEERTKFAVSIVVRSSSDITEPNILSETIEQVSA 311
QY 302 EVLRLEQKASKMKDLVLKKKAELEEHRRRAHLVGEEGYAEFEIATAEAGIDPSLVLE 361
Db 312 EVDVCFNKLKASRMKELVWKRTLENLCLAHIEADTSTSEKSTALIDSLGVDPSLELT 371
QY 362 QIEAHIAITVKEAFSRKDIILEKVERWQVACBEEAWLDYKDDNRYNAGRAHUTLKRAE 421
Db 372 NIELHINKIKBEAHSRKEIIDRIDWLSACEENWLEYNQDETYSAGRGHYNLKHAE 431
QY 422 KARTLVNKKIPGMVDVLRKIAAWKNRGEKEDFTVDGVSLSMLDEYMFVRKEQEKQRQ 481
Db 432 PARITVANKISWVDNLIKKILLWEDET-RKSFYDGVRLVSLIDEDYKLTIRKQGEKKRY 490
QY 482 RDQKLODQKAOEALYSGSPSKPLSTKTK-----APRHSMGGA 522
Db 491 RDQKQMDLLIKRESYSGSPSKPRSSVRKTYNGYNGDASVPTPRNSAGA 543

RESULT 14
Q9LVBI PRELIMINARY; PRT; 511 AA.
AC Q9LVBI;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE EMB|CAB82688.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RL DNA Res. 7:31-63 (2000).
DR EMBL; AB019235; BAA97189.1; -.
DR InterPro; IPR002048; BF-hand.
DR InterPro; IPR007145; MAP65_ASE1.
DR Pfam; PF03999; MAP65_ASE1; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
SQ SEQUENCE 511 AA; 59754 MW; 475B3C0F7FC80E78 CRC64;

Query Match 34.4%; Score 1202; DB 2; Length 511;
Best Local Similarity 46.5%; Pred. No. 9.1e-47;
Matches 244; Conservative 113; Mismatches 134; Indels 34; Gaps 8;

QY 11 QMTTCDLLELVINWDEYGEPTTDRMLLEQCELVYRRKVDQANRSRAQLKRAI 70
Db 5 QIESTWSLLQLEIILWKEVGETETREKILIEIEECREYVNRKIEKVEEKIRKQEI 64
QY 71 ABGEALAGCSAMGPPV---HVRSQNKHLG--LRLEALNAPYLEEMKKKVRWQ 125
Db 65 ADSEARVIDICSVMEBPPIGLRHQSDQSGNRSRLKDELVKILQKLEEKERKSEKIQ 124
QY 126 FVHVIEQIKKISSEIR-----PADFPFKVPVQDQSLRLKLDLTVDLESLOKES 177
Db 125 FIQVDDIRCVREINGESDDETCSDFF-----SADESLSLRKLDLHRELHRELTLQEQKR 179
QY 178 DRUKQVIEHLNLSHLCVGLIDFKQTVYVHPSLDEAGSKNLSNTTIERLAAANRLR 237
Db 180 NRVKIQDNIIRTLLESCLSVLGNFRFETVTKIHPSLVDVTGSGRSISNETLDKASSVQW 239
QY 238 EMKIQRMKQLODFASSMLELWNLMDTPLEEQQMFQNTCNIAASQEITEPNTLSTDFLN 297
Db 240 ETKIORMQELQDLVTTMLFNLMDTPAEBOQKFMVDVSCNIAATVSEITKPNLSLDLLE 299
```

QY 298 VYSEVRLLEQLKAGKQKDLVLKKAELHEHRRRAHLV-GBEGYAEFSIEAIEAGADP 356
Db 300 EVKALCLREELKWKSKMLVLKKSSELEICRTHIVLEEDIAENVVIRKIESGVNP 359
QY 357 SLVLEQIEAHIAIVTKKEAFSRKDIILEKVERWQNACEEAWLEEDYKNDNRYNAGRGHULT 416
Db 360 ENILEQIEVRACKVKEEALSRLKEKADKWLNACEEENWLEEVNQDENRYNAGKGSHLI 419
QY 417 LKRAEKARTLVNKKIPGVVDVLTAKAENKRGKEDFTYDGVSLSSMLDEYMFVRQKEQ 476
Db 420 LKRAEKARLVNKLPAWVEALASKITIWESEK-EYEFUDSEPTNARVYGTQRRRTIR 478
QY 477 -EKKQRQDKLQDLKAEALYGSKPSKPLSTKKAPRHSMG 520
Db 479 TPQKESSETSRSSDIRQRKCY-----EAPKREKG 510

RESULT 15

Q6Z113 PRELIMINARY; PRT; 589 AA.
AC Q6Z113;
DT 05-JUL-2004 (TremBLrel. 27, Created)
DT 05-JUL-2004 (TremBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TremBLrel. 27, Last annotation update)
DE Putative microtubule-associated protein MAP65-1a.
GN Name=OJ1008_D06.15;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP04040; BAD12872.1; -
DR InterPro; IPR007145; MAP65_ASE1.
DR InterPro; IPR009061; Putativ_DNA_bind.
PF Pfam; PF03999; MAP65_ASE1; 1.
SQ SEQUENCE 589 AA; 66465 MW; 42AA084B3C3D54A3 CRC64;

Query Match 33.8%; Score 1182.5; DB 2; Length 589;
Best Local Similarity 42.9%; Pred. No. 8.3e-46;
Matches 255; Conservative 133; Mismatches 176; Indels 31; Gaps 12;
QY 1 MSSAVKDLQHQMSSTTCDSLLELNVIWDEVGEPDTRDRMLLEBOECLEVYRKVDQAN 60
Db 1 MSALLRE-----TSCGSLILKQLQSVNDEVGESEEDRKVLFQDQECLEVDYKRVDOAT 54
QY 61 RSPAQLRKAIAEAGEALAGICSAMGPPVHVRSQNKHLGRLNNAIVPYLEEMKKKV 120
Db 55 KSRDLLQALDYKTELARLASALGKSIDI-SPEKTARTIKQLTAIAPTLEQLGKKK 113
QY 121 ERNMQFVHVEIQIKTSSEIRPADFPFKV--PVDQSLSLRKDLTKDLSLQKES 177
Db 114 ERIKELANTOSRIEQIRGEIAGTLEMGQQVALPQINEDDLTVRKLRFPQLQLQLEKES 173
QY 178 DRLKQVIEHLNLSLCEVLGIDFKQTVVHVPSLDEAGS--KNLSNTTIERLAAAANR 235
Db 174 RLLEKVLHVGVHDLNVLGMDFFRITITQVHSSLDSDSIGNHEKNISNETLSKLDRTIGT 233
QY 236 LREMKIORMKQLODPASSMLELNMDTPLEEQMQFNITCNIAASEQIEPTNTLSTDF 295
Db 234 LNEDKRLRLKLEQLATQLYDLMDLMDTPVEERSLFDHVSNCNRTATVEEVVPGALADV 293
QY 296 LNVVESVRLLEQLKAGKQKDLVLKKAELHEHRRRAHLVGBEGYAEFSIEAIEAGAD 355
Db 294 IDQQTVEERLDQKYSKMKIEAFKKQAILEDIYASTHVVDTAVAHEKIQALESNME 353
QY 356 PSLVLEQIEAHIAIVTKKEAFSRKDIILEKVERWQNACEEAWLEEDYKNDNRYNAGRGHULT 415
Db 354 PSELIADWDSQILKAKEALSRLKEKADKWLNACEEENWLEEVNQDENRYNAGKGSHLI 413

QY 416 TLKRAEKARTLVNKKIPGVVDVLTAKAENKRGKEDFTYDGVSLSSMLDEYMFVRQKE 475
Db 414 NLKRAEKARILVSKIPALVETLVAKTRAWEENHGL-PPMYDGVSLAMLDEYVILRQBRE 472
QY 476 OEKKQRQDKLQDLKAEALYGSKPSKPLSTKK-APRHSMGGA-----NRRLSL 528
Db 473 EEKXRMREQRQTQLLNIDREGPFGRVNVRYVTSAKKVAGTKPENGASNGTSPSRLST 532
QY 529 GGATWQPPKPTDILHSKSV-RAAKKTEIEIGTLPSSSRGLDIAGLPIKKLSFNAST 582
Db 533 G-----NQLNESKSTGRSAGKDKKGA-SKNTATSLNEAA-PADKEAADST 577

Search completed: November 20, 2004, 00:42:10
Job time : 109 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 2004, 00:14:11 ; Search time 25 Seconds

(without alignments)
2655.593 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKDLQHQWTTCDLSL.....DIDYFEERRLAIVLARQMV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79;*

1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1362.5	39.0	648	T47895	hypothetical prote
2	1208	34.6	608	F84430	hypothetical prote
3	1042	29.8	587	E84808	hypothetical prote
4	1004	28.7	473	T04799	hypothetical prote
5	905	25.9	592	E86404	hypothetical prote
6	259.5	7.4	885	S59660	anaphase spindle e
7	230	6.6	880	F75103	conserved hypothet
8	228.5	6.5	1972	A41604	myosin heavy chain
9	222	6.4	1827	T16270	hypothetical prote
10	221	6.3	1780	T17272	hypothetical prote
11	221	6.3	1938	JC5421	smooth muscle myos
12	221	6.3	1972	JC5420	smooth muscle myos
13	217	6.2	2116	A26655	myosin heavy chain
14	214	6.1	886	H69378	conserved hypothet
15	213.5	6.1	1963	1 MKWK	myosin heavy chain
16	211	6.0	1927	A59236	embryonic muscle m
17	210.5	6.0	830	T00029	Miranda protein -
18	210.5	6.0	2020	T21174	hypothetical prote
19	209	6.0	848	A44972	paramyosin - nemat
20	209	6.0	1937	T138055	myosin heavy chain
21	208	5.9	936	T339083	myosin heavy chain
22	208	5.9	1256	T26101	hypothetical prote
23	208	5.9	1690	T13030	microtubule bindin
24	207.5	5.9	876	T23767	myosin heavy chain
25	207	5.9	676	S00084	myosin heavy chain
26	207	5.9	741	T339082	myosin heavy chain
27	207	5.9	2101	T42184	nuclear mitotic ap
28	206.5	5.9	1175	T35815	myosin heavy chain
29	206.5	5.9	1201	A35815	myosin heavy chain

ALIGNMENTS

RESULT 1

T47895 hypothetical protein T4C21.250 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C:Accession: T47895

R.Cholsene, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; S.

.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24479

A:Accession: T47895

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-648 <CHO>

A:Cross-references: UNIPROT:Q9LZY0; EMBL:AL162295

A:Experimental source: cultivar Columbia; BAC clone T4C21

C:Genetics:

A:Map position: 3

A:Introns: 61/3; 145/3; 214/3; 266/3; 320/1; 370/3; 400/1; 424/3; 449/3; 535/1

A:Note: T4C21.250

Query Match	39.08;	Score	1362.5;	DB 2;	Length	648;			
Best Local Similarity	45.1%;	Pred. No.	3.6e-57;						
Matches	307;	Conservative	124;	Mismatches	184;	Indels	65;	Gaps	18;
Qy	30	VGEPTTRDRMLLEQRCLEVYRRKVDQANRSRAQLRKAIAGEAELAGICSAMGEPV	89						
Db	1	MGETEDEKADSLADIEKECLSVYKRKVEASRGKANLLKEIAVGRAEIAAIGSSMGQEI	60						
Qy	90	HVRQSNOKL-HGLREELNAIVPYLEEMKKKKVRRWQFVHVIEQIKKISSEI-PPADFPV	147						
Db	61	H--SNSRGENLKEELENNVQDGLRKRAERMFNEVIDQLKLSLQLGHPNTDYLK	117						
Qy	148	FKVPVDDQSLRLKLDLTKDLESQKESDRLKQVIEHLNSLHSLCEVLGIDPFQTVYE	207						
Db	118	-KFAAEETDLSLQRLLEELRSQGLQNEKSKRLEEVECLLKTLSLCSVLGDFKGMIRG	176						
Qy	208	VHPSLDEAGSKNLSNTTIERLAAAANRLREMKIORMQKLODFASSMLELWNLMDTPLEE	267						
Db	177	IHSSLVDS-NTRDVSRSSTLDKLDMMIVNLRKLRQMKQVQDLAVSLLELWNLDTLPAEE	235						
Qy	268	QOMFONITCNIAAQEITEPNTLTSTDFLNYVESEVLRLEQLKASKMKDILVKKCAELE	327						
Db	236	QKIFHNVTGIALTESEITANILSVASIKRVEDEVIRLSKIKITIKKEVILRKRLELEE	295						
Qy	328	HRRAHLVGEEGYAEFEFTEAIEAGAIIDPSLVLEQIEAHIAIVKEEAFSRKDIILEKVERW	387						
Db	296	ISRKMHMATEVLKSENFSEVATESGVKDPPEQLLEQIDSEIAKVKKEASRSKLEKVEKW	355						
Qy	388	QNACBEEAMLEWYKNDNRYNAGRGHAHLTKRAEKARTLVNKIPGMVDVLTAKIAWQNE	447						
Db	356	MSACBEEWLEENRDNRYNAGRGHAHLTKRAEKARLLVNLKPLGMVEALTAKVTAWNE	415						

Db 423 IPVFQFNKFWALLFSPKQFKFSNIIDYCNSTTHSMVDVLTTKVKAWEKRG-VPF 481
QY 454 TYDGVLSNMLDEYMFVROEKEQKQRQDKLQDLKABQALYGSFSPSKPL----- 509
Db 482 LCDKQPLLOTDDIVIRAQEEKQFQREKRLQGLQGLATEKEAKYSAXKPLGQSL 541
QY 510 ---STKAP--RHSMGAGNRRLSLGG 530
Db 542 NTDNVTKTPIGRIGNTPGRSVTSGG 567

RESULT 4
T04799
hypothetical protein F10M23.100 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C/Accession: T04799
R:Bevan, M.; Lecharny, A.; Chafador, F.; Krivitzky, M.; Kreis, M.; Hoheisel, J.; Mewes, H.
submitted to the Protein Sequence Database, February 1999
A/Reference number: Z15385
A/Accession: T04799
A/Molecule type: DNA
A/Residues: 1-473 <BEV>
A/Cross-references: UNIPROT:Q9SZ16; EMBL:AL035440
A/Experimental source: cultivar Columbia; BAC clone F10M23
C/Genetics:
A/Map position: 4
A/Introns: 24/3; 92/3; 184/3; 256/3; 308/3; 412/3
A/Note: F10M23.100

Query Match 28.7%; Score 1004; DB 2; Length 473;
Best Local Similarity 46.9%; Pred. No. 1.9e-40;
Matches 204; Conservative 94; Mismatches 113; Indels 24; Gaps 4;

QY 15 TCSLLLELVNVIWDEYGEPTTDRMLLEQECLEVYRKYQDQANRSRAQLKATAEGE 74
Db 15 TCGLTQLQKLEIWDVEGSEDEERDKLLQIEBECVLYKKVKKVLAAKSRAELLQTLSDAT 74
QY 75 AELAGICSANGEP-----PVHVRQSNQKLGRLBELNAIVPYLEMKKKKVERWQ 125
Db 75 VELSNLTALGKSYDIDCSMSLFLQPDKTSGTIKEQLSAIPALEQLMOQKEERVRA 134
QY 126 FVHVIEQIKKISSEIR-----PADFVPKVPVDQSLRLKDLTKDLESQKESDR 179
Db 135 FSDVQSQIKICEIAGLNGPH-----VDETDLSLKRLDDPQKLOELQKESDR 187
QY 180 LKQVIEHLSLCEVLGIDFQTVYVHPSLDEAEG--SKNLSNTTIERLAAAANRLR 237
Db 188 LQKLVFVSTVHDLCAVRLDFLSTVEVHPSLDEANGVQTKSINETLARLAKTVTLK 247
QY 238 EMKIQRKQLQDPASSMLELWNLMDTPLEBQQMFQNTCNIAASEQIEITEPTNLSTDFLN 297
Db 248 EDKQRLKQLQELATQLTDLWNLMDTSDERELFDHVTNISASVHEVTASGALALDLE 307
QY 298 YVESEVLRLQKASKWKOLVLKKALEHRRRAHLVGEVGAEEFSAIAGADIPS 357
Db 308 QAEVEVDRLDQLKSSRNKEIAFKQSELEIYARAHIEIKPEVVRERIMSLIDAGNTEPT 367
QY 358 LVLEQIEAHATVKEEAFSRKQILEKVERQWNAEEAWLEDYKNDNRYNAGRGAHLTL 417
Db 368 ELLADMQSIQAKAKEAFSRKEILDRVEKWKMSACEESWLEDYNRDQNRYSRGAHLNL 427
QY 418 KRAEKARTLVNKKIPG 432
Db 428 KRAEKARILVSKITG 442

RESULT 5

E86404
hypothetical protein F13K9.3 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C/Accession: E86404
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federpiel, N.A.; Kaul, S.; White, O.; Alonso, K.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.M.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luro, J.S.; Maity, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712
A/Accession: E86404
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-592 <STO>
A/Cross-references: UNIPROT:Q9C7G0; GB:AE005172; NID:gl1024863; PIDN:AAG26947.1; GSPDB: C/Genetics:
A/Map position: 1

Query Match 25.9%; Score 905; DB 2; Length 592;
Best Local Similarity 38.6%; Pred. No. 1.2e-35;
Matches 210; Conservative 117; Mismatches 195; Indels 22; Gaps 10;

QY 23 LNVIWDEYGEPTTDRMLLEQECLEVYRKYQDQANRSRAQLKATAEGEALAGICS 82
Db 57 LNMWDEYGEDEKFEREKVLLDIEQECVEAYRKYRKHANVSRSLHQELAESEALTHFL 116
QY 83 ANGEPPVHVRQSNQKLGRLBELNAIVPYLEMKKKKVERWQFVHVIEQIKKISSEI-- 140
Db 117 CIGERSVPGR-PEKGGTLREQLDSIAPALREMRUKDERVQKPSVKEIKQKISEIAG 175
QY 141 -RPADFVPKVPVDQSLRLKDLTKDLESQKESDRKQVIEHLSLHSLCEVLGI 199
Db 176 RSTYEDSTFKITIDNDLSNKKLEVEQNELHRLHDEKNERLQKVDIYICARDLSATLGT 235
QY 200 DFKQTVYVHPSLDEAEG--SKNLSNTTIERLAAAANRLREMKIQRKQKLODFASSMLELW 258
Db 236 EASMITTKIHPSLNDLYGISKNISDILKKLNGTVVSLSEEEKHLEKLEKHLHILGRALSNLW 295
QY 259 NIMDTPLEBQQMFQNTCNIAASEQIEITEPTNLSTDFLNVESEVLRLQKASKMKDLV 318
Db 296 NIMDASVDEQRQFHFVIDLSSAPSDVCAPGSITDIIQAAEAEVRLDQLKASRIKELF 355
QY 319 LKKKALEHRRRAHLVGEVGAEEFSAI-EAEGAIDPSLVLEQIEAHATVKEEAFSR 377
Db 356 IKQKELEDTCNWSM--ETPSTENGNTNLVDSGEVDHVDLLAAMDKEIARAKEEAASR 413
QY 378 KDILEKVERQWNAEEAWLEDYKNDNRYNAGRGAHLTLKRAEKARTLVNKKIPGAVDVL 437
Db 414 KGIIEKVDRLMLASDEERWLEEDQDENRYSVSRNAHRLRRAERARITVSKISGLVESI 473
QY 438 RTKIAKWKVERKEDEFTYDGVLSNMLDEYMFVROEKEQKQRQDKLQ--DQLKABOE 496
Db 474 LVKAKSWEVERQKV-FLYNEVPLVAMLQYKNLQKREMEKQKREKKNKSIDPQVAEGD 532
QY 497 ALYGKSPSPKPLSTKKAPRHSNGGA-----NRLSLGGATMQPPKTDILHKSVRAA 549
Db 533 NYFMARPAS-----SNPRISNRSMNGGFGSGSPINRKYSGGFNTNNNTALGTSIRRES 587
QY 550 KXTE 553
Db 588 RKSE 591

RESULT 6
S59660
anaphase spindle elongation protein ASE1 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein O2806; protein YOR058c
C/Species: Saccharomyces cerevisiae
C/Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 09-Jul-2004
C/Accession: S59660; S66941
R:Pellman, D.; Fink, G.R.

submitted to the EMBL Data Library, January 1995
A:Description: Yeast microtubule-associated proteins required for anaphase spindle elong
A:Reference number: S59660
A:Accession: S59660
A:Molecule type: DNA
A:Residues: 1-885 <PEL>
A:Cross-references: UNIPROT:P50275; EMBL:U20235; NID:g972941; PID:AAA75026.1; PID:g9729
R:Bohn, C.; Bolotin-Fukuhara, M.; Daigman-Fornier, B.; Dang, D.V.; Valens, M.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S66929
A:Accession: S66941
A:Molecule type: DNA
A:Residues: 1-885 <BOH>
A:Cross-references: EMBL:Z74966; NID:g1420196; PID:CAA99251.1; PID:e252338; PID:g142019
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD:ASE1
A:Cross-references: SGD:S0005584; MIPS:YOR058C
A:Map position: 15R

Query Match 7.4%; Score 259.5; DB 2; Length 885;
Best Local Similarity 20.7%; Pred. No. 5.8e-05;
Matches 173; Conservative 150; Mismatches 288; Indels 225; Gaps 38;
QY 1 MSSAVKQDLQHSTTCDSLLLELVIVDEVGEPDT---TRDRMLLEQECLIEVYRRKVD 57
DB 86 MDNMYRENFILISKOLEKLENLAVIQNYGYSNTEIITKEKIIFTTISNSIKQFFEQAD 145
QY 58 QANRSRAQRKAIAGEAB---LAGICSAMGE---PPVHVR-----OSN 95
DB 146 E-----ELKRLSAENGIEQDILNNILRINDPSGIKTPIDLYIRNAILLQESKTVPOSP 199
QY 96 QKHLGLREE-----LNAIVPYLEEMKKKVERWQVHVIEIOIKKISSIRPAD 144
DB 200 KXPLSLKKAALDAKAFVLGSPFLRLDYLSUI---TLKHIIQSKENLPGCLTEAD 255
QY 145 -----FVPFKVPVDQS---DLSRLKLDLTKDL-----ESL----- 172
DB 256 NEAIAEFPELSTLTAYLIQENGKGDIGLSMKFIIDNRKDIILKGSFAFTINEESVKHNE 315
QY 173 -----QKESDLKQVIEHLNSLHSLCEVLGT-----DPKQ-----TVY 206
DB 316 VIKYIEEYERRFKSVLTCKVSISSICQLGTPLATLIGDFEQLRSYGEENSTSEIP 375
QY 207 EYHPSLDAEGSKNLSTNTTIERLAAANRLREMKIORMKQDQFASSMLELWNLMDTLE 266
DB 376 NFHP-VDRERNK--IDITLEKQAIHKERADKKLLMEQCQ-----KLWTRLKISQE 425
QY 267 EQOMFONITCNIAASEQITEPTNTLSTDFLNYVESEVLRLQLKASKMKDLV---LKKA 323
DB 426 YIKTF-----MRNNSLSTESLGRISKVRMLEAMKKLIKLLISDWDKIQ 472
QY 324 EL-----EHRRAHLVGEYAEESIEAIEAGADPSLVLEQIEAHATVKEEAFS 376
DB 473 ELWRTLQVSEERSKFIIVFEELRNSATLQ-----EDELLETCENELKLEEKLT 525
QY 377 RKDILEKVERWONACEEBAWLEDYNDKDNRYNAGRAHLLKRAEKARTLVNK-IPGMVD 435
DB 526 YKPIKLISDFESLQDOEFLERSKDSRL-LSNSHKILLTEEMKWKIRTRHPPRVIN. 584
QY 436 VLRTKIAWKNRGKED---FTYDGVSLSSMLDEYVFRQEKQEKORQDKLQDLKA 493
DB 585 DLRIKL---EADGLFDQPFLLFKGKPLSEAD-----IQQGEIAKYPRCVRMRQSKG 636
QY 494 EQEALYKSPSPKPLSTKAPRHSWGMGANNRLSLGGA--TWQPP-----KTILH 542
DB 637 KCGANKENKVIKNTPKATESIRVDIG-----LNLDANITYKTPSKKTIQGLTKNDLSQ 691
QY 543 SKSV-RAAKTTEEIGT-----LSPS-----SSRGLDIAGLPIKLSFNASTLRETETP 589
DB 692 ENSLARHMQGTIKLSPNRRATRLAAPTIVSRNSKG-----NIERPILNRN--RSSDLS 743
QY 590 KXPAQITPG-NSVSVSTPVRPTNTNTEDDEN-----RTPKFTALNPKTPMT-----VTA 638

DB 744 SSPRINTHGEHAKVPRQLPPIPLNKVDTKGSHIPQLTKKALELLKRESTGTCENVR 803
QY 639 PMQAMTPTSLANKVSATPSLVYDKPEVTL-----QEDIDYSPEE 678
DB 804 PERKSSLEDYAQKLSSP-----YKEPESHYIKLSMSPGKFLQNIQKDIESGDD 854
RESULT 7
F75103
conserved hypothetical protein PAB0812 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F75103
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: F75103
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-880 <KAW>
A:Cross-references: UNIPROT:Q9U2C8; GB:A0248286; GB:AL096836; NID:g5458366; PIDN:CAB5013
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB0812
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF1032
Query Match 6.6%; Score 230; DB 2; Length 880;
Best Local Similarity 21.1%; Pred. No. 0.0014;
Matches 148; Conservative 121; Mismatches 228; Indels 206; Gaps 32;
QY 41 LLEEQECLEVYRKV---DQANRSRAQLKATKAEGAEAGELAGCSAMGE---PPVHVQRS 94
DB 292 LOEKEE-----YRKIKGFRDEYKLRLEKLSKWESELKAEVIEKGEKKKERABEI 347
QY 95 NOKLGLRLEELNATVPLYEEMK-----KKVERW-----NQFVHVIEQIKKISSEI 140
DB 348 REKLSEIEKRLLEELKPYVEELEDKQVQKQIERLKLKGLSPGEVIEKLESLEKETE 407
QY 141 RPADFVFPKVPVDQSDLSRLKDLTKDLESLOKESDRKLQVIEHLNSLHSLCEVLGID 200
DB 408 EEA-----IKEITTRIGQMEQEKNERMK-AIEELRKAKGKCPVCGRE 448
QY 201 F-----KOTVVEVHPSLDEAGSKNLSTNTTIERLAAAN-ELREMKIORMKQDQFASSML 255
DB 449 LTEHKKELMERY-TLEIKIEELKKTTEERKLRLVNLKRLHKLKREFSVMRDIAEQIK 507
QY 256 ELNMLMDTPLEEQOMFONITCNIAASEQITEPTNTLSTDFLNYVESEVLRLQ-LKASK- 313
DB 508 EL-----ESKLK-----GFNLEELQEKEREPEGLNEEF-NKLKGELGLERDLKRIKA 554
QY 314 -----MKDLVLKKALEEHRRRAHLVCEGY-----AEPFSIAIEAGADIPS 357
DB 555 LEGRRLKIEKVRKAKAELENLHRLQELGFEVSEELNLRIQBLEEPHDKYVEAKKSESE 614
QY 358 LVLEQIEAHATVKEEAFSRKDLKVKVERWQACEE-EAWLEDYNDKDNRYNAGRAHLLT 416
DB 615 --LRELKNLEKTELDQAFEMLADVE---NIEEKEAKDL---ESKFNE----- 659
QY 417 LKRAEKARTLVNKPIMGVVLRTKIAAWKNRGKEDFTYDGVSLSSMLDPMFVROKQEQ 476
DB 660 -EYVEEKRLRLVLEKREVSLSLTARLEELK-----SVEQIKATLRKLKEEKEE 706
QY 477 EKKRQDKKLQDLKAEQALYGSKPSPSKPLSTKAPRHSWGMGANNRLSLGATWQPP 536
DB 707 REKAEIKKLEKAL----- 721
QY 537 KTDILSKSVRAAKTTEEIGTLSPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQI 596
DB 722 -----SKVEDRKKIKDYKTLA-----KEQALN-----RISEIASEIFSEF 757
QY 597 TPGNSVSVSTPVRPTNTNTEDDENRTPKFTALNPK-TPMTVTA-----PMQAMTTP 646

Db 758 TDG-KYSNVIVRA-----ENKTKLFVVYEGKEVPTLTLGGGRIALGLAFRLALSM 808
QY 647 SLANKVSVATPSVLVDYKDPVETVQEDIDYSEFERRLAIVARQM 689
Db 809 YLVGRIDL----LILDEPTPFLDEE-----RRRLKLDIMERHL 842

RESULT 8
A1604
myosin heavy chain, smooth muscle, long splice form - rabbit
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 19-Apr-2002
C:Accession: A41604; A43501
R:Babji, P.; Kelly, C.; Periasamy, M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10676-10680, 1991
A>Title: Characterization of a mammalian smooth muscle myosin heavy-chain gene: complete
A:Reference number: A41604; MUID:92073350; PMID:1961735
A:Accession: A41604
A:Molecule type: mRNA
A:Residues: 1-1972 <BAB>
A:Cross-references: GB:M77812
R:Nagai, R.; Kuro-o, M.; Babji, P.; Periasamy, M.
J. Biol. Chem. 264, 9734-9737, 1989
A>Title: Identification of two types of smooth muscle myosin heavy chain isoforms by cDN
A:Reference number: A33501; MUID:89255535; PMID:2722872
A:Accession: A33501
A:Molecule type: mRNA
A:Residues: 1455-1972 <NAG>
A:Cross-references: GB:J04833; NID:gi65519; PIDN:AAA31407.1; PID:gi65520
A:Experimental source: smooth muscle
A>Note: examination by Southern blotting for the regions of difference between this isofo
active splicing
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; alternative splicing; ATP; coiled coil; hydrolase; methylated
F:88-771/Domain: myosin motor domain homology <MMOT>
F:178-185/Region: nucleotide-binding motif A (P-loop)
F:559-572/Region: actin binding #status predicted
F:633-647/Region: actin binding #status predicted
F:844-1938/Domain: coiled coil #status predicted <COI>
F:844-1284/Region: S2
F:1285-1972/Region: light meromyosin
F:1939-1972/Domain: carboxyl-terminal <CBT>
F:129/Modified site: N6,N6-Crithylllysine (Lys) #status predicted
F:184/Binding site: ATP (Lys) #status predicted
F:701,711/Active site: Cys #status predicted

Query Match 6.5%; Score 228.5; DB 1; Length 1972;
Best Local Similarity 19.6%; Pred. No. 0.0048;
Matches 151; Conservative 159; Mismatches 316; Indels 143; Gaps 30;

QY 4 AVKDQLHQ-MSTTCDSLLL-----ELNVIVDSVGPDTTDRMLLELEQ---ECLEVR 53
Db 1150 ALKTEEDTLDTATQQLRAKREQEVTVLLKALDESTRSHEAQVQEMRKHTQVVEELT 1209
QY 54 RYVDQANRSRAQL----RKATAEAGELAGICSANGPPVHVR----- 92
Db 1210 EQLQEFKRAKLANLDKTKQTLKENADLAGSLRVGQAQVEHKKKLEVLQELQSKCS 1269
QY 93 -----QSNOKLHGLREELNAIVPYLEEMKKKVERWQFVHVIEQIKKISSEIRPADF 145
Db 1270 DGRARAEKNDKVKHQKNEVSTGMLSEAGKAIK-----LAKEVASLGSQQLQDTQE 1322
QY 146 VPFKVPVQDQLSLRKLDLTDLKQLESQKESDRL--KQVIE-HLNSHLCEVLGIDFX 202
Db 1323 LLQETRTQLNVS-TKLQLEDERNLSQELDEMEAKQNLERHISTLN---IQLSDSK 1377
QY 203 QTVEVHPSLDEAGSKNLSNTTIERL-----AAANRLREMKQRMOKLQDPASSML 255
Db 1378 KKLQDPASTVESLEEGKKRFQKTESLTQOYEEKAAAYDKLEKTKNRLQQLDLDLVVDLD 1437
QY 256 ELNWLMDTPLEEQOMF-----QNITCNIA-----ASEQEITEPTNTLSTDFLNVSEVL 304

Db 1438 NORQLVSNLEKKQKKFDQLLAEKKNISSKYADERDRAEAAREKETKALSARALEEAL 1497
QY 305 RLQQL-KASKMKDLVLKKAELEHRRRAHLVGBEGYAEFEFSIEAIEAGAI DPSLVLEQI 363
Db 1498 AKHEELRTNKM-----LKAEMEDLVSSKDDVGKVVHELEKSKRALETQMEEMKQLEEL 1551
QY 364 EAHATVKEEAFSRKMDILEKVERMQNACEEAWLEWYKNDNRYNAGRAH---LTLKRA 420
Db 1552 EDELQAT-EDAKLRLEV--NMQALKVQFERDLQARDEQNEEKRRQLQRLQHEYELEDE 1608
QY 421 EKARTLV-----NKTPGMVDVLRTKIAWKNVERGKEDFTYDGVLSLMDLEYMFVQE--- 473
Db 1609 RKQRALAAAKKLEGG--DLKDLQLQADSAIKGEEAIKQLLKQAQMKQFORELEDA 1666
QY 474 -----KQOEKKR---ORDOKQLQDLKAEQALYGS---KPSPSKPLSTKKAPHS 518
Db 1667 SRDEIFATAKENEKAKSLEADLMQLQEDLAAERAPKQADLEKELAEELASSLSGRNA 1726
QY 519 MGGANRRRLSLGGATMQ---PPKTDILHSKSVRAAKKTBEEIGTLPSSSSRGLDIAGLPICK 575
Db 1727 LQDEKRRLEARIQAULEEEEBEQNMEASDRVRKATQQAELS-----NE 1772
QY 576 LSFNASTLRETETPRKPAQITPGNSVSTVPRPTNNNTEDDENRTPKTFTALNPKTPT 635
Db 1773 LATERSTAQKNESARQ---QLERQNKELSKLQEMEGAV---KSKPKSTIALEAK----- 1822
QY 636 VTAPWQMAWTPSLANKVSAATPSVLVDK--PVTLQEDIDYSEFERRLA 682
Db 1823 -IAQLEEQVEQAREKQAAAKALKORDKKLEMLLOVE-----DERKMA 1865

RESULT 9
T16270
hypothetical protein F35D11.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16270
R:Fulton, B.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F35D11.
A:Reference number: Z18487
A:Accession: T16270
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1827 <FUL>
A:Cross-references: UNIPROT:Q20042; EMBL:U29391; NID:g868214; PID:g868224; PIDN:AAA6875
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CBSP:F35D11.11
A:Introns: 76/2; 131/3; 159/3; 185/3; 221/3; 253/3; 320/1; 869/3; 1133/3; 1205/2; 1250/3

Query Match 6.4%; Score 222; DB 2; Length 1827;
Best Local Similarity 18.5%; Pred. No. 0.0088;
Matches 192; Conservative 143; Mismatches 289; Indels 414; Gaps 37;

QY 4 AVKDQLHOMSTTCDSLLLELVIVDSVGPDTTDRMLLELE-----QECLEV---YRR 54
Db 492 SLKGDVAAMKTDLDKTLRLDML-----ETERDELKKKLETEREQADQORDLEIAECRA 544
QY 55 KYVDQANRSRAQLRAKIAEGEA----- 75
Db 545 KLDEMAEKEALRKELAEAFQAITAMEGEGKLNQEQFLESKNKELNLTLDQIESLNSEVEN 604
QY 76 ---ELAGICSANGPPVH---VRQSNOKLHGLREELNAIVPYLEEMKKKVERWQFVHV 129
Db 605 KNEIRNLMTLQEKVEHVIQNVRTSSHQLTATYEEANGSIDILKA-----ELTRL 654
QY 130 IEQIKKISSEIRPADFPVPKVPVQDQLSLRKLDLTDLKQLESQKESDRLKQVIEHLNS 189
Db 655 HEQVNRNTRKQISEAN-----EKYDDARKNDALLEDVATWQ-EKYEQKMLESEMNR 705
QY 190 LHSCLCEVLGIDFKQTVYEVHPSLDEAGSKNLSNTTIERLAAANRLREM----- 239

Db 706 RQEKEREADRLALLDDLGRNFDKLTNELKQGVTVDSLNEBIEISSLEQNLKSKERKE 765
 Qy 240 KIQRMQKL-QDPASSMELWNL-MDTPLLEEQQFONITCNIAASEQET-----EPN 289
 Db 766 ELLRMEELFGKNEAEKMEYEVKQLQAEKQGVENFQKECEARNELTKIHEMLMEHD 825
 Qy 290 TLSTDFNLVSEVLRL-----OLK 310
 Db 826 QLKVDHL-HTEVERLEKMKRKELEKLEKNEQDGRPAEWSNERNRLESSKNEAVTELQER 884
 Qy 311 ASKMKDLVLKXK-----AELEHRRRAHLVGE----- 338
 Db 885 VQLEDDVVKEDKEIALRRDLEDGSHKRDLDLKLKRWELTDESKEDRKEQKTLNEE 944
 Qy 339 -----GYAF-EFSIAEAGADP 356
 Db 945 RMLMEQKEEAMLVATKHATTIDQOTRRISVLGDEVEKLTAGIAERESSINALESNWEL 1004
 Qy 357 SLVLEQIEAHIAIVKKEAFSRKDILEKVERWQACE--EEAWLEDYNK-----DNRYNAG 410
 Db 1005 ISKLETTAELEKJDEL---AVMLKQNSLKNKGEGSEKWEERKKIQDLADQLREAN 1061
 Qy 411 RGAH-----LT----- 416
 Db 1062 KVVHNMKNVNLBEKKNELQNVDTLTKVQRLEIQLMDKAAKNEVSGDLLRKMEHDAQ 1121
 Qy 417 -LKA-----EKAR-TLVNKIPGMVDVLTAKA--KNERKEDFTYDGVSL 461
 Db 1122 SMLKQAEQEFRLTLEKVRKALODENORLVNDLATVKAAPFVKRETSK-----AIS 1174
 Qy 462 SMLDEYMFVRQEKQEKQRQDKKLQD-----LKAEQALYSGKSPSKPLSTPKA-- 514
 Db 1175 DILDY---RSABEKANGELDNQRLSDLATVTLKLRQEL-YAKDSNRLRDSQKPE 1230
 Qy 515 -----PHSMGGANRL-----SLGGATMQPPKTTDILHKSVR 547
 Db 1231 EVQSKLANLQKSAVESLQNPMSNRQNSRYVDIPRAASISGLNENSDVPLRSSPSVR 1290
 Qy 548 RAKKTEE---IGTSLSPSSRGLDIAGLP--IKKLSFNASTLRET-----ETPRKPPAQI 596
 Db 1291 FADSQNNQRAVDSMDVSSGVTVLFLKERIEQLEADNADLSDALEKAKDELQRNEKL 1350
 Qy 597 TPGNSVSTPVPRITNTEDD---ENRTPKFTALNPKTPTMTVAPMQAMT----- 645
 Db 1351 ADROMVIERVERQLVHITEENTIENR-----MTSQRMVLTNEESSRSR 1395
 Qy 646 -----PSLANKVSATPVSLVYDKPE 665
 Db 1396 EHEIRSMKARISTLEHLREKESKLAHLRKEIEVLHGQLHDALESKEKATGLVGVQDSKH 1455
 Qy 666 VTLEQDIDYSPFEERLAI 683
 Db 1456 RDLEEQLDRANERELAI 1473

RESULT 10
 T17272
 hypothetical protein DKFp434B0435.1 - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Oct-1999 #sequence revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T17272; T46451; A36881
 R:Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, September 1999
 A:Reference number: Z18723
 A:Accession: T17272
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1780 <POU>
 A:Cross-references: UNIPROT:O9URF5; EMBL:AL117496
 A:Experimental source: adult testis; clone DKFp434B0435
 R:Angorge, W.; Wirker, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23028
 A:Accession: T46451
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 575-1136, 'C', 1138-1608, 'K', 1610-1780 <AAA>
 A:Cross-references: EMBL:AL137392
 A:Experimental source: adult testis; clone DKFp434I152
 R:Westendorf, J.M.; Rao, P.N.; Gerace, L.
 Proc. Natl. Acad. Sci. U.S.A. 91, 714-718, 1994
 A:Title: Cloning of cDNAs for M-phase phosphoproteins recognized by the MP2 monoclonal
 A:Reference number: A36881; MUID:94119956; PMID:8290587
 A:Accession: A36881
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1215-1261, 'E', 1263-1418, 'S', 1420-1608, 'K', 1610-1780 <WES>
 A:Cross-references: GB:116782; NID:G292328; PIDN:AAC37542.1; PID:G292329
 C:Genetics:
 A:Note: DKFp434B0435.1; DKFp434I152.1
 C:Keywords: phosphoprotein

Query Match 6.3%; Score 221; DB 2; Length 1780;
 Best Local Similarity 20.98; Pred. No. 0.0095;
 Matches 164; Conservative 124; Mismatches 305; Indels 190; Gaps 33;
 Qy 1 MSSAVKQDLHQMSTTCDLSLLELNVIMDEVGE---PDTTRDRMLLEQECLEVYRRKV- 56
 Db 995 VSKQVRYIQEPNRENSPHSSIEAIWEECKEIVKASSKSHQIEELEQQ-IEKLAQAEVK 1053
 Qy 57 ---DOANRSRAQLKATAEAEALAGICSANGEPVHVVRQSNQKLHGLRBLNAIVPYLE 113
 Db 1054 GYKDENNRLK-----EKE-----HKNQDDL---LKEKETLIQQLKE 1086
 Qy 114 EMKKKKYERNQFVHVIEQIKKISSEIRPADFPVFPVQSDLSLRKLD-----ELTKD 168
 Db 1087 ELQEKNTLDVQHQHVEGKRAUSELTQGTVCYAKIKELETILETKQVSHSAKLEQD 1146
 Qy 169 LESIQKE---KSDR-LKQVIEHLNLSLCEVLGIDFKQTVVYVHPFSLDEAGSKNLSN 223
 Db 1147 I--LEKESIILKLERNLKEFOEHLQ-----DSVKNTKDLNV 1180
 Qy 224 TTI---ERLAAANRLBEMKIQRMKLODPASSMELWNLMDTPLEEQQFONITCNIAA 280
 Db 1181 KELKKEEITQLTNNLQDMKHLQLKEEBEETNRQTEKLELSASSARTQNLKADLQR 1240
 Qy 281 SEQEITPNTLSDF---LNVSESEVLRL---EOLKASKMKDLVLK---KAELEEHRRR 331
 Db 1241 KEEDYADLKEKLTDAKQIKQVQKQSVWDEDEKLLRIKINELEKKKQCSQELDMKQRT 1300
 Qy 332 AHLVGE--GYAEFFSIEAIEAGIDPSLVLEQIEAHIAIATVKEEAFSRKD-----ILE 382
 Db 1301 IQQLKEQLNNQKVEEAIQOYERACKDLNVKEKIIDMRMTLEEQEQTVQEQDQVLEAKLE 1360
 Qy 383 KVERWQACBEEAWLEDYN---KDDNRYNAGRAHLLTKRAEKARTLVNKKIPGMVDVLR 438
 Db 1361 EVERL--ATELEKWKECNDLETNNQNSN-----KEHENNTDVLGKLTNLQDELQ 1409
 Qy 439 TKIAAWKNRGEKEDFTYDGVSLSDMLDEYMFVRQEKQEKRRQDOKKLOD-----OLK 492
 Db 1410 ESEQYNADRK-----WLEKMWLITQAEAEINRNKMKYAEADREFFKQ 1458
 Qy 493 AEQALYSGKSPSKPLSTTKAPRHSNGGA-----NRRLSLGATWQPP 536
 Db 1459 NEMEILTALQTERDSDILQKWEERDQLVAALQLKALISSNVQKQNEIEQLKRIISETS 1518
 Qy 537 K--TDILHKSVRRAK-----KTEEIGTSLSPSSRGLDIAGLPKIKLSFNASTLRETET 588
 Db 1519 KIETQIMDIKPKRISSADPKLOTEPLSTSEFSRNKIEDGSSVVLDSCVESTENDQSTRF 1578
 Qy 589 PRKP-----EAQITPGNSVSTP--VRPIT-----NNTEDD-----ENRTPK 623
 Db 1579 P-KPELEIQITPLQPNKMAVKHFGCTTPVTVEIPKARKRKSNEEEDLVCKENKQATPR 1637
 Qy 624 TTTALNPKTPM-----TVTAPMQAMTP-----SLANKVSATPVSLVYDKPEVTLEQD 671

Db 362 -RPKMDRMGQIKAKLEKULTDPKVEKMDYLLSKAKKEBEKTEKLLKLIAKKSSLKTRG 420
Qy 259 NMDPTPLEQQMFPNTCHIAASEQEITPNTLSDFP---LNVSEVLRQLQ----- 309
Db 421 AQLKKAVERKLSAER-TCVCGRELDEEHRKIMAEYTRMKRIABELAKADEIKKLE 479
Qy 310 -----KASKMKDVLKXK-----AELEHRR-RAHLVGEQY 340
Db 480 RLEKVEKALEKETVLKPYQWDELKALENELSSHDAEKLSAESEYKVKERLDGLRQ 539
Qy 341 ABEFSTEATEAGAPDPSL-----VLEQIEAHIA-----TVKEAFSKOILEKVERQWAC 391
Db 540 QKILLSSASRIKELSSLRLEIBALKNVESERGELHRIKREGEF---ESLEELER----- 591
Qy 392 EEEAMLEDYKNDNRNAGRGHLLTKRAEKARTLVNKPIMGVMDVLTRTIAAKNERGKE 451
Db 592 EVQSLRPFYNNKMLEKDAESRLESELKRRK-----LEDEISEIAKLEANGKA 641
Qy 452 DFTYDGVSLSSLMDEYMFVRQEKQEKQRQDKQLQD--QLKAEQALYGGKPSKPL 509
Db 642 E-----ETRGQIDELLRIYSEEHRLSDEHLRKSKELAGLSRLETRESLSQSAEKOL 695

RESULT 15

MWKw

Myosin heavy chain B [similarity] - Caenorhabditis elegans

N;Contains: myosin Arpase (EC 3.6.4.1)

C;Species: Caenorhabditis elegans

C;Date: 13-Jun-1983 #sequence revision 19-May-2000 #text change 09-Jul-2004

C;Accession: T20770; T21629; A93958; A93287; A21074; A02592

R;Kershaw, J.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19322

A;Accession: T20770

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1963 <WIL>

A;Cross-references: UNIPROT:O02244; EMBL:Z81499; PIDN:CAB04089.1; GSPDB:GN00019; CESP:F1

A;Experimental source: clone F11C3

A;Accession: T21629

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-1963 <W12>

A;Cross-references: EMBL:Z83107; PIDN:CAB05505.1; GSPDB:GN00019; CESP:F11C3.3

A;Experimental source: clone F32A7

R;Karn, J.; Brenner, S.; Barnett, L.

Proc. Natl. Acad. Sci. U.S.A. 80, 4253-4257, 1983

A;Title: Protein structural domains in the Caenorhabditis elegans unc-54 myosin heavy ch

A;Reference number: A93958; MUID:83273600; PMID:6576334

A;Accession: A93958

A;Molecule type: DNA

A;Residues: 1-61, 'EMSVIQ', 65-376, 'V', 378-1963 <KAR>

A;Cross-references: GB:J01050; NID:g156399; PIDN:AAA28124.1; PID:g156400

R;McLachlan, A.D.; Karn, J.

Nature 299, 226-231, 1982

A;Title: Periodic charge distributions in the myosin rod amino acid sequence match cross

A;Reference number: A93287; MUID:82272395; PMID:7202124

A;Accession: A93287

A;Molecule type: DNA

A;Residues: 847-1333, 'R', 1335-1876, 'L', 1878-1963 <MCL>

R;Wills, N.; Gesteland, R.F.; Karn, J.; Barnett, L.; Bolten, S.; Waterston, R.H.

Cell 33, 575-583, 1983

A;Title: The genes sup-7 X and sup-5 III of Caenorhabditis elegans suppress amber nonsen

A;Reference number: A21074; MUID:8322892; PMID:6571695

A;Accession: A21074

A;Molecule type: DNA

A;Residues: 1873-1963 <W13>

A;Cross-references: GB:V01494; GB:J01049; NID:g6783; PIDN:CAA24738.1; PID:g6784

C;Genetics:

A;Gene: unc-54; CESP:F11C3.3

A;Map position: 1

A;Introns: 21/3; 64/3; 111/3; 264/1; 525/3; 951/2; 1747/3; 1819/3; 1894/3

C;Superfamily: myosin heavy chain; myosin motor domain homology
C;Keywords: actin binding; ATP; coiled coil; hydrolase; methylated amino acid; muscle c
F;84-775/Domain: myosin motor domain homology <MMOT>
F;174-181/Region: nucleotide-binding motif A (P-loop)
F;662-684/Region: actin binding #status predicted
F;766-780/Region: actin binding #status predicted
F;848-1963/Domain: coiled coil #status predicted <COI>
F;848-1162/Region: S2
F;1163-1963/Region: light meromyosin
F;125/Modified site: N6,N6,trimethyllysine (Lys) #status predicted
F;180/Binding site: ATP (Lys) #status predicted
F;702,712/Active site: Cys #status predicted

Query Match 6.1%; Score 213.5; DB 1; Length 1963;
Best Local Similarity 20.6%; Pred. No. 0.024;
Matches 150; Conservative 125; Mismatches 295; Indels 157; Gaps 27;

Qy 1 MSSAVKQDLQHQSTTCDSLLLELVINWDEGEPTDTRMLLEBOECLEEVYRRKVDQAN 60
Db 1313 LKSQTSQLEEARRTADEEARQRTVAAQAKNYQHEAQQLQESLEE-IEGNEILRQLS 1371
Qy 61 RSRAQLR--KATAEGEAEAGICSAMGPPVHVRSQNKHLGLREELNAIVPYLEMKKK 118
Db 1372 KANADIQWKARFEGE---GLLKA-DELEDAKRQAQKINELQALDAANSKNASLEKT 1426
Qy 119 K-----VERWNQFVHVIEQ---IKKISSEIRPADFPVFKVPVQDSLSLRK 151
Db 1427 KSLRGVGLDDAQVDVERANGVASALEKKQKQKGIIDWR-----KK 1468
Qy 162 LDELTKDLESLOKESDRLLKQVIEHLNLSHLSCEVL-GI-----DFKQTVYVHPSLDEA 215
Db 1469 TDDLAELDGAQRDLRNTSTDLFKAKNAQAEELAEVVEGVRRENKSLUSQBIKDLTDQLEG 1528
Qy 216 EGSKNLSNTTIERLAAANRLREMQRKQKQKQDFASSMLEL--WNLMPTPEEQQMFQN 273
Db 1529 GRSVHEMKIIRRL-----EIEKELOHALDEAAEALEAESKVLRAQVEVSQIRSE 1580
Qy 274 ITCNTAASQETETPNTLSTDFLNVSSEVLRLQKAS-----KMDLVLLKKA--- 323
Db 1581 IEKRIQKEEE-----FENTRKNHARALESMQASLETEAKGKAEILLRIKKLEGD 1630
Qy 324 ---ELE---EHRRAHLVGEEGVAAEFSEIEAIEAGALDPSLVLEQIEAHATVKEEAFS-- 376
Db 1631 INELEIHALDHANKADAKQ-----NLKRYEQVRELQVLEEEORNGADTRQGFNAE 1684
Qy 377 -----RKDIL---EKVERWQNAEEAEWLEDYKDDNRVYAGRGHLLTKRAEKAR 424
Db 1685 KRATLQSEKELLVANAEEARQAQAEYA--ADARQANEANAQVSSILTSKR----- 1737
Qy 425 TLVNKIPGMVDVLRTKIAAAWKNR-RGKEDFTYDGVSLSSMLDEYMFVRQEKQKQRD 483
Db 1738 ---KLEGEIQAIHADLDLTLNEYKAAEERSKKAATADATRLABEL--RQEQEHSQHVDR 1791
Qy 484 QKKLQDLK-----AQEALYSGKPSKPSPLSTTKAPRHSMGGANRL-----SLGG 530
Db 1792 RKGLEQKLEIQVRLDEAEAAALKGGKVIKLEQVRVLESELDGEQRRFQDANKNLG- 1850
Qy 531 ATMQPPTDILHKSVRRAAKTKEEIGTLSPSSSRGLDIAGLPKIKLSFNASTLRETETPR 590
Db 1851 -----RADRRVRELQFQVDEKKNFERQLDLIDKLQ-----QKLTQK 1888
Qy 591 KPFAQITPGNSVSTPVRPITNNEDENRTPKTFTALNPKTPMTWTAPMQMWTSLAN 650
Db 1889 KQVEEAELANLNLQYKQLTHQLEDAERADQ---AENSLSKVRSKSRASASVAPGLQS 1945
Qy 651 KVSATPV 657
Db 1946 SASAAVI 1952

Search completed: November 20, 2004, 00:42:40

Job time : 31 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 19, 2004, 19:16:05 ; Search time 87 Seconds

(without alignments)
2845.094 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKQDLHQNSITTCDSL.....DIDYFERRLAIVLARQMV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesep 238Sep04:*

1: Genesep1980s:*

2: Genesep1990s:*

3: Genesep2000s:*

4: Genesep2001s:*

5: Genesep2002s:*

6: Genesep2003as:*

7: Genesep2003bs:*

8: Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3490	99.8	690	5	Abb05052 Rice leaf
2	3477.5	99.5	690	5	Aau97649 Plant sha
3	1236.5	35.4	578	3	Aag20828 Arabidops
4	756.5	21.6	338	3	Aag20829 Arabidops
5	693.5	19.8	318	3	Aag20830 Arabidops
6	325.5	9.3	620	6	Abu70658 Human adi
7	323.5	9.3	620	8	Adn04802 Antipsori
8	323.5	9.3	620	8	Adn04802 Antipsori
9	323.5	9.3	620	8	Adn04802 Antipsori
10	319.5	9.1	657	4	Abd11530 Human cyt
11	313.5	9.0	670	7	Adb78256 Human CGD
12	267.5	7.7	639	8	Adp04701 Sea squir
13	240	6.9	1857	5	Aau84350 Protein M
14	240	6.9	1938	6	Abn92127 Human cer
15	240	6.9	1945	6	Abu10398 Human smo
16	240	6.9	1972	6	Abn92126 Human cer
17	240	6.9	1972	7	Adn58724 Human Pro
18	240	6.9	1972	7	Adn58721 Human Pro
19	240	6.9	1972	7	Adn58720 Human Pro
20	240	6.9	1972	7	Adn58720 Human Pro
21	240	6.9	1972	7	Adn58720 Human Pro
22	240	6.9	1972	7	Adn58720 Human Pro
23	240	6.9	1972	7	Adn58720 Human Pro
24	237.5	6.8	953	4	Abu53071 Intracell
25	237.5	6.8	2228	7	Abn61599 Human gol

26	237.5	6.8	2230	6	ABU07445	Abu07445 Protein d
27	237.5	6.8	2230	7	ABR61600	Abn61600 Human gol
28	237.5	6.8	2250	7	ABR61601	Abn61601 Human gol
29	237.5	6.8	2252	7	ABR61602	Abn61602 Human gol
30	237	6.8	941	4	ABU53072	Abu53072 Intracell
31	237	6.8	953	4	ABU53069	Abu53069 Intracell
32	235	6.7	860	5	AD117302	Ad117302 Polypepti
33	235	6.7	860	7	ADC38517	Adc38517 Human AML
34	233.5	6.7	1455	7	ADN80783	Adn80783 Microsate
35	230	6.6	880	4	ABN96332	Abn96332 Putative
36	229	6.6	885	2	AAE66930	Aae66930 AMLL chro
37	229	6.6	885	2	AAE66930	Aae66930 AMLL chro
38	221.5	6.4	907	4	ABU53073	Abu53073 Intracell
39	221.5	6.3	876	7	ADM26084	Adm26084 Hyperthar
40	221	6.3	951	4	ABU53070	Abu53070 Intracell
41	221	6.3	1114	7	ADJ70313	Adj70313 Human hea
42	221	6.3	1213	4	AAW40016	Aaw40016 Human pol
43	221	6.3	1213	6	ABP97031	Abp97031 Human L-F
44	221	6.3	1780	4	AAW38681	Aaw38681 Human pol
45	221	6.3	1780	4	ABU53201	Abu53201 Human cel
					AAW25750	Aaw25750 Human pro

ALIGNMENTS

RESULT 1
ABB05052
ID ABB05052 standard; protein; 690 AA.
XX
AC ABB05052;
XX
DT 26-MAR-2002 (first entry)
XX
DE Rice leaf shape and size controlling protein SEQ ID NO:2.
XX
KW Rice; leaf shape and size controlling gene; leaf shape; leaf size.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT Misc-difference 170
FT /label= unknown
FT /note= "encoded by GNA"
XX
PN JP2001258574-A.
XX
PD 25-SEP-2001.
XX
PF 23-MAR-2000; 2000JP-00083067.
XX
PR 23-MAR-2000; 2000JP-00083067.
XX
PA (NORQ) NORINSUISANSHO NOGYO SEIBUTSU SHIGEN.
PA (SEIB-) SEIBUTSUKETI TOKUTEI SANGYO GIJUTSU.
PA (NORI-) ZH NORIN SUISAN SENTAN GIJUTSU SANGYO.
XX
WP1: 2002-037587/05.
DR N-PSDB; ABA92702.
XX
PS A new gene controlling leaf shape.
XX
PS Claim 1; Page 9-11; 14pp; Japanese.
XX
CC The present sequence represents the protein encoded by a polynucleotide
CC (see ABA92702) encoding a plant gene which can control leaf shape and
CC size. The polynucleotide can be used for controlling leaf shape
XX
SQ Sequence 690 AA;

Query Match 99.8%; Score 3490; DB 5; Length 690;
Best Local Similarity 99.9%; Pred. No. 1.7e-245;
Matches 689; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSSAVKQDLQHQMTTCDLSLLLELNVWDEVGPDPTTRDRMLLEBQECLEVVYRRKVDQAN 60
Db |
QY 1 MSSAVKQDLQHQMTTCDLSLLLELNVWDEVGPDPTTRDRMLLEBQECLEVVYRRKVDQAN 60
Db |
QY 61 RSRAQLRKAIAEGEAEAGLCSAMGEPVHVVRQSNQKHLGLREELNAIVPYLEEMKKVKV 120
Db |
QY 61 RSRAQLRKAIAEGEAEAGLCSAMGEPVHVVRQSNQKHLGLREELNAIVPYLEEMKKVKV 120
Db |
QY 121 ERWNOFVHVHIOIKKISSEIRPADVFPFKVPVDSLSLRKLDLTKDLESLOKESDRL 180
Db |
QY 121 ERWNOFVHVHIOIKKISSEIRPADVFPFKVPVDSLSLRKLDLTKDLESLOKESDRL 180
Db |
QY 181 KQVIEHLNLSHLSLCEVLGIDFKQTVYVHPSDEAEGSKNLSNTTIERLAAAANRLREMK 240
Db |
QY 181 KQVIEHLNLSHLSLCEVLGIDFKQTVYVHPSDEAEGSKNLSNTTIERLAAAANRLREMK 240
Db |
QY 241 IORMKQLQDFASSMLELWNLMDTPLEEQMFQNTICNIAAQEITEPNTLSTDFLNVE 300
Db |
QY 241 IORMKQLQDFASSMLELWNLMDTPLEEQMFQNTICNIAAQEITEPNTLSTDFLNVE 300
Db |
QY 301 SEVLRLQOLKASKMKDLVLLKKAELREHRRRAHLVGESEYAEFSEIAEAGIDPSLV 360
Db |
QY 301 SEVLRLQOLKASKMKDLVLLKKAELREHRRRAHLVGESEYAEFSEIAEAGIDPSLV 360
Db |
QY 361 EQIEAHIAIVKEAFSRKDIKLEKVERWQACEEAEWLEDYNDKDDNRYNAGRAHILTKRA 420
Db |
QY 361 EQIEAHIAIVKEAFSRKDIKLEKVERWQACEEAEWLEDYNDKDDNRYNAGRAHILTKRA 420
Db |
QY 421 EKARTLVNKPIMGVMDVLRKIAAMKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQK 480
Db |
QY 421 EKARTLVNKPIMGVMDVLRKIAAMKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQK 480
Db |
QY 481 QRDOKKLQOLKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db |
QY 481 QRDOKKLQOLKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db |
QY 541 LHSKSVRAAKTTEIGTGLSPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Db |
QY 541 LHSKSVRAAKTTEIGTGLSPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Db |
QY 601 SVSSTPVRPITNNTEDDENRTPKTFTALNPKTMTVTAPMQAMTPSLANKVSATPVSLV 660
Db |
QY 601 SVSSTPVRPITNNTEDDENRTPKTFTALNPKTMTVTAPMQAMTPSLANKVSATPVSLV 660
Db |
QY 661 YDKPEVTLOEDIDYSFEERLAIYLARQMV 690
Db |
QY 661 YDKPEVTLOEDIDYSFEERLAIYLARQMV 690
Db |

RESULT 2
AAU97649
ID AAU97649 standard; protein; 690 AA.
XX
AC AAU97649;
XX
DT 13-AUG-2002 (first entry)
XX
DE Plant shape regulating protein.
XX
KW Plant; shape; rice.
XX
OS Oryza sativa.
XX
FH Key Location/Qualifiers
FT Misc-difference 690 /note= "Encoded by TAA"
FT
XX
PN JP2002125675-A.
XX
PD 08-MAY-2002.
XX
PF 19-OCT-2000; 2000JP-00320111.
XX

PR 19-OCT-2000; 2000JP-00320111.
XX (DOKU-) DOKURITSU GYOSEI HOJIN NOGYO SEIBUTSU SH.
XX
DR WPI; 2002-449190/48.
XX N-PSDB; ABK52591.
PT A new plant gene, modification of a plant by using the gene, and a plant
PT body prepared by the method, a polynucleotide, a vector, and a plant
XX body.
PS Claim 1; Fig 12; 20pp; Japanese.
XX
XX This invention relates to the polynucleotide and protein sequences of a
XX plant shape regulating protein. The invention also comprises a vector
XX containing a controlling sequence or the nucleotide sequence of the
XX invention connected operably to the controlling sequence, a vector
XX containing the antisense sequence of the plant shape regulating gene and
XX a method for modifying a plant. This method includes introducing the
XX vector to a plant tissue to give a transformant, a step of regenerating
XX the transformant to give a plant body and a step of selecting the plant
XX body on the desired character, and a plant body transformed by the above
XX vector. The gene is used for providing a transgenic plant having various
XX characters. The present sequence represents the plant shape regulating
XX protein of the invention
SQ Sequence 690 AA;
Query Match 99.5%; Score 3477.5; DB 5; Length 690;
Best Local Similarity 99.7%; Pred. No. 1.4e-244;
Matches 688; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 MSSAVKQDLQHQMTTCDLSLLLELNVWDEVGPDPTTRDRMLLEBQECLEVVYRRKVDQAN 60
Db |
QY 1 MSSAVKQDLQHQMTTCDLSLLLELNVWDEVGPDPTTRDRMLLEBQECLEVVYRRKVDQAN 60
Db |
QY 61 RSRAQLRKAIAEGEAEAGLCSAMGEPVHVVRQSNQKHLGLREELNAIVPYLEEMKKVKV 120
Db |
QY 61 RSRAQLRKAIAEGEAEAGLCSAMGEPVHVVRQSNQKHLGLREELNAIVPYLEEMKKVKV 120
Db |
QY 121 ERWNOFVHVHIOIKKISSEIRPADVFPFKVPVDSLSLRKLDLTKDLESLOKESDRL 180
Db |
QY 121 ERWNOFVHVHIOIKKISSEIRPADVFPFKVPVDSLSLRKLDLTKDLESLOKESDRL 180
Db |
QY 181 KQVIEHLNLSHLSLCEVLGIDFKQTVYVHPSDEAEGSKNLSNTTIERLAAAANRLREMK 240
Db |
QY 181 KQVIEHLNLSHLSLCEVLGIDFKQTVYVHPSDEAEGSKNLSNTTIERLAAAANRLREMK 240
Db |
QY 241 IORMKQLQDFASSMLELWNLMDTPLEEQMFQNTICNIAAQEITEPNTLSTDFLNVE 300
Db |
QY 241 IORMKQLQDFASSMLELWNLMDTPLEEQMFQNTICNIAAQEITEPNTLSTDFLNVE 300
Db |
QY 301 SEVLRLQOLKASKMKDLVLLKKAELREHRRRAHLVGESEYAEFSEIAEAGIDPSLV 360
Db |
QY 301 SEVLRLQOLKASKMKDLVLLKKAELREHRRRAHLVGESEYAEFSEIAEAGIDPSLV 360
Db |
QY 361 EQIEAHIAIVKEAFSRKDIKLEKVERWQACEEAEWLEDYNDKDDNRYNAGRAHILTKRA 420
Db |
QY 361 EQIEAHIAIVKEAFSRKDIKLEKVERWQACEEAEWLEDYNDKDDNRYNAGRAHILTKRA 420
Db |
QY 421 EKARTLVNKPIMGVMDVLRKIAAMKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQK 480
Db |
QY 421 EKARTLVNKPIMGVMDVLRKIAAMKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQK 480
Db |
QY 481 QRDOKKLQOLKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db |
QY 481 QRDOKKLQOLKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db |
QY 541 LHSKSVRAAKTTEIGTGLSPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Db |
QY 541 LHSKSVRAAKTTEIGTGLSPSSSRGLDIAGLPIKLSFNASTLRETETPRKPPAQITPGN 600
Db |
QY 601 SVSSTPVRPITNNTEDDENRTPKTFTALNPKTMTVTAPMQAMTPSLANKVSATPVSLV 660
Db |

Db 600 SVSSTVETITNNEDENRKTFTALNPKPTMTVTADPMQMTPLANKVTSATPVSIV 659
 Qy 661 YDKPEVTLOEDIDYSEERLAIYLARQMV 690
 Db 660 YDKPEVTLOEDIDYSEERLAIYLARQMV 689

RESULT 3

AG20828
 ID AAG20828 standard; protein; 578 AA.

AC AAG20828;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 23163.

KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.

XX Arabidopsis thaliana.

XX EP1033405-A2..

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.
 PR 05-MAR-1999; 99US-0123180P.
 PR 09-MAR-1999; 99US-0123548P.
 PR 23-MAR-1999; 99US-0125788P.
 PR 25-MAR-1999; 99US-0126264P.
 PR 29-MAR-1999; 99US-0126785P.
 PR 01-APR-1999; 99US-0127462P.
 PR 06-APR-1999; 99US-0128234P.
 PR 08-APR-1999; 99US-0128714P.
 PR 16-APR-1999; 99US-0129845P.
 PR 19-APR-1999; 99US-0130077P.
 PR 21-APR-1999; 99US-0130449P.
 PR 23-APR-1999; 99US-0130510P.
 PR 28-APR-1999; 99US-0131449P.
 PR 30-APR-1999; 99US-0132048P.
 PR 30-APR-1999; 99US-0132407P.
 PR 04-MAY-1999; 99US-0132484P.
 PR 05-MAY-1999; 99US-0132485P.
 PR 06-MAY-1999; 99US-0132486P.
 PR 06-MAY-1999; 99US-0132487P.
 PR 07-MAY-1999; 99US-0132863P.
 PR 11-MAY-1999; 99US-0134256P.
 PR 14-MAY-1999; 99US-0134218P.
 PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134321P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.

PR 16-JUN-1999; 99US-0139452P.
 PR 16-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-0142920P.
 PR 12-JUL-1999; 99US-0142977P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.

PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137750P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138840P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144332P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145194P.
PR 22-JUL-1999; 99US-0145195P.
PR 23-JUL-1999; 99US-0145221P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 11-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149428P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.

```

PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 21-OCT-1999; 99US-0160980P.
PR 21-OCT-1999; 99US-0160981P.
PR 21-OCT-1999; 99US-0160988P.
PR 21-OCT-1999; 99US-0161404P.
PR 21-OCT-1999; 99US-0161405P.
PR 21-OCT-1999; 99US-0161406P.
PR 21-OCT-1999; 99US-0161359P.
PR 21-OCT-1999; 99US-0161360P.
PR 21-OCT-1999; 99US-0161361P.
PR 21-OCT-1999; 99US-0161920P.
PR 21-OCT-1999; 99US-0161922P.
PR 21-OCT-1999; 99US-0161933P.
PR 21-OCT-1999; 99US-0162142P.

Query Match      21.6%; Score 756.5; DB 3; Length 338;
Best Local Similarity 45.6%; Pred. No. 1.3e-46;
Matches 154; Conservative 69; Mismatches 104; Indels 11; Gaps 4;

QY 241 IORMKQLQFASSMELWNLMDTPEEQOMFONITCNIAASQOITEPTNTLSTDFLNYVE 300
Db 1 MQRLLQLQELATQLTDLNLMMDTSEERELFDHVTNSISASVHVETASGALALDLIEQAE 60

QY 301 SEVLRLQKASKMKDLVLKKAELHRRRAHLVGEEGYAEFEISIEAEGAI DPSLVL 360
Db 61 VEVDRLDQLKSRMKEIAFKQSELEIYARAHIBIKPEVVREIRIMSLIDAGNTEPTTELL 120

QY 361 EQIEAHIAIVKEASRSKDI LKVERWQACEEAWLEDYKDDNRYNAGRAHILTKRA 420
Db 121 ADMDSQIAKAEAFSRKEILDRVEKWSACEESWLEDYKDDNRYNAGRAHILTKRA 180

QY 421 EKARTLVNKPQWVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKKR 480
Db 181 EKARILVSKITAMVDTLAKTAWEEENSMSFEYDGVPLLAMLDYETMLRQERDEKRR 239

QY 481 QRDQKKLQDLKAEQALYGVSKPSKPLSTKKAPRHSWGGANRLSLGGATMQPPKTDI 540
Db 240 LKEQKQKQBPHTQDESAPGSPAPVSAKKP-----VGRVNGGGLNETPMRRLS 292

QY 541 LHSKSVRAAKTEBIGTL-SPSS--SRGLDIAGLPIKK 575
Db 293 MNSNQNGSKRDSLNKIASPSNIVANTKDDAASPVR 330

```

```

RESULT 5
AAG20830
ID AAG20830 standard; protein; 318 AA.
XX
AC AAG20830;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 23165.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.

```

```

PR 05-MAR-1999; 99US-01231180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126284P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.

```

PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151910P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.

PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 19.8%; Score 693.5; DB 3; Length 318;

Best Local Similarity 45.3%; Pred. No. 4.8e-42; Mismatches 101; Indels 11; Gaps 4;
Matches 144; Conservative 62;

QY 261 MDTPLEEQMFONITCNIAASQEQITEPTNTLTSTDFLNYVESEVLRLEQLKASKMKDLVLK 320
DB 1 MDTSEERELEFDHVTNSISASVHVETASGALALDLIEQAEVEVDRLDQLKSSRMKEIAFK 60
QY 321 KKAELERHRRRAHLVGEEGYAEFEISIEAIEAGAI DPLSVLEQIEAHIAITVKEEAFSRKDI 380
DB 61 KQSELEETVARAHIEIKPEVVRERIMSLIDAGNTEPTTELLADMDSDQIAKAKKEEAFSRKEI 120
QY 381 LEKVERWQACEEAEWLEDYKDDNRYNAGRAHLTLKBAEKARTLVNKI PGWVDVLRTK 440
DB 121 LDRVEKMSACESEWLEDYKDDNRYNAGRAHLTLKBAEKARTLVNKI PGWVDVLRTK 180
QY 441 IAAWKNERGKEDFTDGVSLSSMLDEYMFVROEKEQKQRDRDQKQLQDLKAEQBALYG 500
DB 181 TRAWEEENSM-SFEYDGVPLLAHLDEYTLWROEREDEKRLKEQKQSQPHPTDQSSAFG 239
QY 501 SKPSPKPLSTKAPRHSNGGANRLSLGGATWQPPKTDILHKSVRAAKKEETIEITL-S 559
DB 240 SKPSPARPVSAKKP-----VGTRVNGGGLNETPMRRLSMNSNQNGSKSRKDSLNIAS 292
QY 560 PSS--SRGLDIAGLPKK 575
DB 293 PSNIVANTKDDAASPVS 310

RESULT 6

ABU70658
ID ABU70658 standard; protein; 620 AA.

XX AC ABU70658;

XX DT 10-JUN-2003 (first entry)

XX DE Human adipocyte Selected Interacting domain, SID, #289.

XX

Query Match	9.3%;	Score 323.5;	DB 8;	Length 620;
Best Local Similarity	22.7%;	Pred. No. 1.2e-14;		
Matches 144;	Conservative 124;	Mismatches 266;	Indels 101;	Gaps 24;
QY	9	LHQMSTTC-DSLLLELVNIVDEVEGPDTRDRMLLELEQCELEVVYRRKVDQANRSRAQLR	67	
Db	7	LAEEISVCLQKALNHLREIWEIGIPEDQRLQRTVEVKKHKEILLDMIAEESLERLI	66	
QY	68	KAIAGEAEALAGICSAMGPPVHVRSQKLGHLREELNAIVPYLEEMKKKKVRRNQFV	127	
Db	67	KSISVCQKELNLTCLSELHVEFPQ-EEGETTILQLEKDLRTQV---ELMRKQKKEKQELK	122	
QY	128	HVIEQIKKISSIRPADFPVFPVVDQSD-LSRLKDELTKLESLOKESDLKQVI--	184	
Db	123	LLQEQDQELCEILCMPHY-----DIDSASVPSLEELNQPHQVTTURETKASREBFVSI	177	
QY	185	---EHLNLSHSLCEVLGIDFKQTVYVHPVPSLDEAGSKNLSNTTIERLAAANRLRMKIQ	242	
Db	178	KRQIILCMBELDHTPDTSFERDV-----VCEDEDAFCLS---LENIATLQKLLRQLEMQ	228	
QY	243	RMQK---LQDFASSMLELNLMDTPLEEQMFQNTICNIAAEEQETEPNTLSTDFLNVY	299	
Db	229	KSQNEAVCEGLRTQIRELWDRLOIPEEREAVATI---MSGSKAKVRKA-----L	275	
QY	300	ESEVLRLQKASKMDLVKKKAEI-----EHRRAHLVGEEGYAEFSEIAI	349	
Db	276	QLEVDRLLEELKQNMKKVIEAIRVELVQWDQCFYSQEQRAFAPCAEDYTES-----	329	
QY	350	EAGAIIDPSVLQIEAHIAITVKEAFSRKIDLEKVERWQACBEEAW-----LEDYKDD	404	
Db	330	-----LLQLHDAEIVRLKNYVEVHKELFEGVQKW-----EETWRLFLFEKASDP	375	
QY	405	NRYNAGRGHILTKRAEKARTLVNKPIMVVDVLRTKIAAWKNGRGEDEFTYDGVLSLML	464	
Db	376	NRF-TNRGGNL-LKEEKQAKLQKMLPKLEELKARIELWEQHSKA-FWVNGQKFM---	429	
QY	465	DEYMFVROKEQ-EKKRQDQKLODQKAEQALYGSKP-----SPSKPLSTKKA	514	
Db	430	-EVASQWHRLEKRAQERQKQKQTEMLYGSAPRTSKRGLAPNTPGKARKL	488	
QY	515	PHRSMGKANRLSL---GGATMQPPKTDILHSKSVRAAKTTEIGTLPSSSRGLDIAG	570	
Db	489	NTTMSNATANSIRPIFGTVVYHSPVSRLLPSSGSKPVAASTCS-GKKTPTGR---HG	543	
QY	571	LPKIKLSFNASTLRETETPRKPPAQITPGNSVST	605	
Db	544	ANKENLELNGSILSGGYGSGAPLQRFNSINSAVST	578	
RESULT 8				
ADO20037				
ID	ADO20037	standard; protein; 620 AA.		
AC	ADO20037;			
XX				
XX	12-AUG-2004	(first entry)		
DT				
DE				
XX	Human PRO polypeptide #473.			
XX	Human; PRO; immune related disorder; systemic lupus erythematosus;			
KW	rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;			
KW	systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;			
KW	autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;			
KW	diabetes mellitus; renal disease; demyelinating disease;			
KW	central nervous system; peripheral nervous system;			
KW	demyelinating polyneuropathy; Guillain-Barre syndrome;			
KW	chronic inflammatory demyelinating polyneuropathy.			
OS	Homo sapiens.			
XX				
XX	WO2004043361-A2.			
PN				
XX				

PD	27-MAY-2004.
XX	
PF	06-NOV-2003; 2003WO-US035268.
XX	
PR	08-NOV-2002; 2002US-0425235P.
XX	(GETH) GENENTECH INC.
PA	Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI	Wood WI, Wu TD;
XX	WPI: 2004-420067/39.
DR	N-PSDB; ADO20036.
XX	
PT	Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT	treating an immune related disorder such as systemic lupus erythematosus,
PT	rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT	spondyloarthropathy.
XX	
XX	Claim 7; SEQ ID NO 946; 173lpp; English.
XX	
CC	The invention relates to human PRO polypeptides and the polynucleotides
CC	encoding them. The polypeptides and polynucleotides are useful for
CC	treating and diagnosing immune related disorders in mammals. The immune
CC	related disorders include systemic lupus erythematosus, rheumatoid
CC	arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC	sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC	haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC	mellitus, immune-mediated renal disease, demyelinating diseases of the
CC	central or peripheral nervous system, demyelinating polyneuropathy,
CC	Guillain-Barre syndrome and chronic inflammatory demyelinating
CC	polyneuropathy. This sequence represents a human PRO polypeptide of the
XX	invention.
XX	
SQ	Sequence 620 AA;
	Query Match
	Best Local Similarity 9.3%; Score 323.5; DB 8; Length 620;
	Matches 144; Conservative 124; Mismatches 266; Indels 101; Gaps 24;
QY	9 LHQMSTTC-DSLLLELVNIVDEVEGPDTRDRMLLELEQCELEVVYRRKVDQANRSRAQLR 67
Db	7 LAEEISVCLQKALNHLREIWEIGIPEDQRLQRTVEVKKHKEILLDMIAEESLERLI 66
QY	68 KAIAGEAEALAGICSAMGPPVHVRSQKLGHLREELNAIVPYLEEMKKKKVRRNQFV 127
Db	67 KSISVCQKELNLTCLSELHVEFPQ-EEGETTILQLEKDLRTQV---ELMRKQKKEKQELK 122
QY	128 HVIEQIKKISSIRPADFPVFPVVDQSD-LSRLKDELTKLESLOKESDLKQVI-- 184
Db	123 LLQEQDQELCEILCMPHY-----DIDSASVPSLEELNQPHQVTTURETKASREBFVSI 177
QY	185 ---EHLNLSHSLCEVLGIDFKQTVYVHPVPSLDEAGSKNLSNTTIERLAAANRLRMKIQ 242
Db	178 KRQIILCMBELDHTPDTSFERDV-----VCEDEDAFCLS---LENIATLQKLLRQLEMQ 228
QY	243 RMQK---LQDFASSMLELNLMDTPLEEQMFQNTICNIAAEEQETEPNTLSTDFLNVY 299
Db	229 KSQNEAVCEGLRTQIRELWDRLOIPEEREAVATI---MSGSKAKVRKA-----L 275
QY	300 ESEVLRLQKASKMDLVKKKAEI-----EHRRAHLVGEEGYAEFSEIAI 349
Db	276 QLEVDRLLEELKQNMKKVIEAIRVELVQWDQCFYSQEQRAFAPCAEDYTES----- 329
QY	350 EAGAIIDPSVLQIEAHIAITVKEAFSRKIDLEKVERWQACBEEAW-----LEDYKDD 404
Db	330 -----LLQLHDAEIVRLKNYVEVHKELFEGVQKW-----EETWRLFLFEKASDP 375
QY	405 NRYNAGRGHILTKRAEKARTLVNKPIMVVDVLRTKIAAWKNGRGEDEFTYDGVLSLML 464
Db	376 NRF-TNRGGNL-LKEEKQAKLQKMLPKLEELKARIELWEQHSKA-FWVNGQKFM--- 429
QY	465 DEYMFVROKEQ-EKKRQDQKLODQKAEQALYGSKP-----SPSKPLSTKKA 514
XX	

```

Db 430 -EYVAEQWEMHRLKXERAKQERQKXKQKTEMLYGSAPRTPSKRRGLAPWTPCKARKL 488
QY 515 PRHSGGANRRLSL-----GGATMQPPKTDILHKSVAARKTEETGTLPSSSRGLDIAG 570
Db 489 NTTMTSNATANGSIRPIFGTGVYHSPVSRPLPPSGSKPVAASTCS-GKKTPTGR----HG 543
QY 571 LPIKLSFNASTLRETEPRKPFQAQITPGNSVSST 605
Db 544 ANKENLELNGSLSGYPGSAPLQRFNSINSVAST 578

RESULT 9
ID ADQ09246
AD ADQ09246 standard; protein; 620 AA.
XX ADQ09246;
XX ADQ09246;
XX 23-SEP-2004 (first entry)
XX Human PRC1 protein SEQ ID NO:431.
DE
XX thanatos-associated protein; THAP; THAP responsive gene; THAP family;
KW THAP responsive element; angiogenesis; inflammation; metastasis; cancer;
KW apoptosis; cardiovascular disease; neurodegenerative disease; chemokine;
KW antiangiogenic; antiinflammatory; cardiovascular; cytostatic;
KW neuroprotective; osteopathic; THAP modulator; THAP synthesis modulator;
KW human.
XX
XX Homo sapiens.
XX OS
XX WO2004055050-A2.
XX
XX 01-JUL-2004.
XX
XX 10-DEC-2003; 2003WO-IB006434.
XX
XX 10-DEC-2002; 2002US-0432699P.
XX
XX 03-JUL-2003; 2003US-0485027P.
XX
XX (ENDO-) ENDOCENT SAS.
XX
XX (CNRS ) CNRS CENT NAT RECH SCI.
XX
XX Girard J, Amalric F, Roussigne M, Clouaire T;
XX
XX WPI; 2004-525034/50.
XX
XX N-PSDB; ADQ09247.
XX
XX Modulating expression of a Thanatos (death)-Associated Protein (THAP)
XX responsive gene for preventing or treating e.g. cancer or inflammation,
XX comprises modulating the interaction of a THAP polypeptide with a nucleic
XX acid.
XX
XX Example 47; SEQ ID NO 431; 612pp; English.
XX
XX The present invention describes a method for modulating the expression of
XX a thanatos (death)-associated protein (THAP) responsive gene. The method
XX comprises modulating the interaction of a THAP-family polypeptide or its
XX biological fragment with a nucleic acid, and so enhancing or repressing
XX the expression of the THAP responsive gene. Also described: (1) a method
XX of modulating the expression of a gene responsive to a THAP/chemokine
XX complex; (2) a pharmaceutical composition comprising a THAP responsive
XX element in a pharmaceutical carrier; (3) a transcription factor decoy
XX consisting essentially of a THAP responsive element; (4) a cell
XX comprising a transcription factor decoy described above; (5) methods of
XX modulating the interaction between a nucleic acid and a THAP-family
XX polypeptide or its biological fragment, or a nucleic acid and a
XX THAP/chemokine complex; (6) a vector packaging cell line comprising a
XX cell comprising a viral vector which comprises a promoter operably linked
XX to a nucleic acid encoding a THAP-family polypeptide or its biological
XX fragment; (7) a method of constructing a cell which expresses a
XX recombinant THAP-family polypeptide; (8) a method of ameliorating
XX symptoms associated with a condition mediated by a THAP/chemokine complex

```

```

CC ; (9) methods of identifying a test compound that modulates transcription
CC at a THAP responsive element or that modulates the transport of a
CC chemokine into the nucleus; (10) methods for reducing the symptoms
CC associated with a condition selected from excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease and neurodegenerative
CC diseases; symptoms associated with a condition resulting from the
CC activity of a chemokine or a THAP-family polypeptide in an individual; or
CC symptoms associated with transcriptional repression or activation
CC mediated by a THAP-family polypeptide in an individual; (11) a vector
CC comprising a THAP responsive promoter operably linked to a nucleic acid
CC encoding a detectable product; (12) a genetically engineered cell
CC comprising the vector described above or that expresses a THAP-family
CC polypeptide or its biological fragment; (13) an in vitro transcription
CC reaction comprising a nucleic acid comprising a THAP responsive promoter,
CC ribonucleotides and an RNA polymerase; and (14) an isolated mutant THAP-
CC family polypeptide that does not bind to a chemokine. The pharmaceutical
CC composition has antiangiogenic, antiinflammatory, cardiovascular,
CC cytostatic, neuroprotective and osteopathic activities, and can be used
CC as a THAP and THAP synthesis modulator. The composition can be used for
CC modulating the expression of a THAP responsive gene. Modulation is useful
CC for reducing symptoms of conditions such as excessive or insufficient
CC angiogenesis, inflammation, metastasis of a cancerous tissue, excessive
CC or insufficient apoptosis, cardiovascular disease or neurodegenerative
CC diseases. The present sequence is used in the exemplification of the
CC present invention.
XX
XX Sequence 620 AA;

```

```

Query Match          9.3%; Score 323.5; DB 8; Length 620;
Best Local Similarity 22.7%; Pred. No. 1.2e-14;
Matches 144; Conservative 124; Mismatches 266; Indels 101; Gaps 24;

QY 9 LHMSTTC-DLLLELVINWDEVEGPDPTDRMLLELEQECLEVYRRKVDQANRSRQLR 67
Db 7 LAESIVCLOKALNHLREIWEIIGIPEDQRLQRTVEVYVKGHIKELDWMIAEESLKERLI 66
QY 68 KAIAGEAELAGICSAMGEPPVHVRSQKQLHGLEELNAIVPYLEEMKKVKVNRQFV 127
Db 67 KSIVCQKELNTLCSELHVFPFO-EGETTILQLEKDLRTQV---ELMRKQKERRKQELK 122
QY 128 HVIEQIKKISSEIRPADFPVPKVPVQSDI--SLRKLDELTKDLESQKESDRUKQVI-- 184
Db 123 LLQEQDQELCELCPHYPH----DIDSASVPSLEELNQFRQHVTTLRKASREEFVSI 177
QY 185 --EHLNLSLCEVLGIDFKQTVYEVHPSLDEAGSKNLSNTTIERLAAAANRLREMKIQ 242
Db 178 KRQIILQWELDHTPDTSPERDV-----VCEDSDAFCLS---LENTATLQKLLRQEMQ 228
QY 243 RMQK---LQDPASSMLLELWMLDTPLEEQMFQNTCNIAASEOEITEPTNLSTDFLNV 299
Db 229 KSNQEAECGELRTQIRELMDRLQIPEERENAVTI---MSGSKAKVKA-----L 275
QY 300 ESEVLRLEQLKASRKMDLVLLKKKAL-----BEHRRRAHLVGBEGYAEBSFIEAI 349
Db 276 QLEVDRLLEELKQNMKKVIEAIRVELVQYDQCFVSEQRQAPAFPFAEDYTES----- 329
QY 350 EAGADISLVLEQTEAHIAITVKEAFSRKQILEKVERWQNCAREAW-----LEDYKDD 404
Db 330 -----LQLQHDAAEIVRLKNKYVEHKLFEVQKWM-----BETWRLFLFEPERKASDP 375
QY 405 NRYNAGRGALHTLRAEKAKRTLNVNKIPQWVDVLRTKIAAWKNGRKGEDFTYDGVLSLSSML 464
Db 376 NRF-TNNGGNL-LKEEKQRAKLQKMLPLSEELKARIELWEQEHSKA-FWVNGQKFM---- 429
QY 465 DEYMFVRQEKEQ-EKKQRQDKKLQDQKAFQEAALYGSKP-----SPSKPLSTKKA 514
Db 430 -EYVAEQWEMHRLKXERAKQERQKXKQKTEMLYGSAPRTPSKRRGLAPWTPCKARKL 488
QY 515 PRHSGGANRRLSL-----GGATMQPPKTDILHKSVAARKTEETGTLPSSSRGLDIAG 570
Db 489 NTTMTSNATANGSIRPIFGTGVYHSPVSRPLPPSGSKPVAASTCS-GKKTPTGR----HG 543

```


AC	ADD78256;	Best Local Similarity 22.5%; Pred. No. 7e-14;	
XX		Matches 156; Conservative 122; Mismatches 262; Indels 153; Gaps 30;	
DT	29-JAN-2004 (first entry)		
XX	Human CGDD-37.		
DE			
XX	Anabolic; Hypertensive; Respiratory; Anti-HIV; Antiallergic;		
KW	Neuroprotective; Nootropic; Antianemic; Antiartherosclerotic;		
KW	Antiinflammatory; Ophthalmological; Muscular; Hepatotropic;		
KW	Neuroprotective; Antiasthmatic; Anticonvulsant; Virucide; Antibacterial;		
KW	Fungicide; Antiparasitic; Protozoicide; Antihelminthic; Cytostatic;		
KW	Cerebroprotective; Antiparkinsonian; Antipsoriatic; Antigout;		
KW	Antidiabetic; Antiarthritic; Antirheumatic; Osteopathic; Gene therapy;		
KW	human; cell growth; cell differentiation; cell death; CGDD;		
KW	cell proliferative disorder; cancer; developmental disorder;		
KW	neurological disorder; autoimmune disorder; inflammatory disorder;		
XX	infection; reproductive disorder.		
OS	Homo sapiens.		
XX			
PN	WO2003077875-A2.		
XX			
PD	25-SEP-2003.		
XX			
XX	14-MAR-2003; 2003WO-US008310.		
XX			
PR	15-MAR-2002; 2002US-0364494P.		
PR	29-MAR-2002; 2002US-0369129P.		
PR	12-APR-2002; 2002US-0372511P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
XX			
PI	Kable AE, Tran UX, Hafalia AJA, Burford N, Honchell CD;		
PI	Lehr-Mason PM, Duggan BM, Ramkumar J, Griffin JA, Richardson TW;		
PI	Elliot VS, Jiang X, Jackson AA, Marquis JP, Chawla NK, Khare R;		
PI	Becha SD, Lee SY, Swarnakar A, Yue H, Warren BA, Baughn MR, Lal PG;		
PI	Lee S, Ho A, Gandhi AR, Yao MG;		
XX			
DR	WPI: 2003-779081/73.		
DR	N-PSDB; ADD78295.		
XX			
PT	New polypeptides and polynucleotides associated with cell growth,		
PT	differentiation and death, useful for diagnosing, treating or preventing		
PT	e.g. developmental, neurological, autoimmune, inflammatory or		
PT	reproductive disorders.		
XX			
PS	Claim 1; SEQ ID NO 37; 320pp; English.		
XX			
CC	The present invention relates to novel human proteins (I; ADD78220-		
CC	ADD78256) and their coding sequences (II; ADD78259-ADD78297), which are		
CC	associated with cell growth, differentiation and death, referred to as		
CC	CGDD-n proteins, where n is a number from 1 to 39. The CGDD proteins and		
CC	their coding sequences are useful for diagnosing, treating or preventing		
CC	cell proliferative disorders (e.g. cirrhosis, hepatitis,		
CC	arteriosclerosis, psoriasis, primary thrombocytopenia) or cancers (e.g.		
CC	adenocarcinoma, sarcoma or cancers of the bone, bone marrow, brain,		
CC	breast, colon, kidney, liver, lung or uterus), developmental disorders		
CC	(e.g. renal tubular acidosis, Becker muscular dystrophy, gonadal		
CC	dysgenesis, hypothyroidism or seizures), neurological disorders (e.g.		
CC	Pick's disease, cataract, epilepsy, ischemic cerebrovascular disease,		
CC	stroke, Alzheimer's disease, Parkinson's disease or dementia),		
CC	autoimmune/inflammatory disorders (e.g. AIDS, allergies, anemia, asthma,		
CC	diabetes mellitus, bronchitis, osteoporosis, osteoarthritis, rheumatoid		
CC	arthritis, contact dermatitis or gout), viral, bacterial, fungal,		
CC	parasitic, protozoan or helminthic infections, reproductive disorders		
CC	(e.g. infertility, ectopic pregnancy, premature ovarian failure, delayed		
CC	puberty or prostatitis) or disorders of the placenta (e.g. preclampsia,		
CC	choriocarcinoma, placenta previa, placental or maternal floor infarction		
CC	or chronic villitis).		
XX			
SQ	Sequence 670 AA;		
XX			
Query Match	9.0%; Score 313.5; DB 7; Length 670;		

QY	9 LHQMSTTC-DSLLLELVNINWDEVEGPDTRDRMLLELEOECELVYRRKVDQANRSRAQLR	67
DB	7 LAEBSIVCLOKALNHLREINELIGIPEDQRLQRTVEVVKKHIKELDDMMMAEESLKERLI	66
QY	68 KAIAGEAEALAGICSAMGEPVHVQRQSNOKLHGLREELNAIVPYLEEMKKKKVERWNOFV	127
DB	67 KSIISVCQKELNTLSELHVEFFQ-RQLAVLS--RQECNGV--SAHCNURLPGSSDFP	120
QY	128 HVIQIKKISSIEIRPADFPVF--KVPVDQSLSLRKLDELTKDLESLOKEKS	177
DB	121 AVASRVAGITG-VRRHAWLIFLVKTGFHHVQGEGETTILQLEKDLRTQVLMRKQKK	179
QY	178 DRLQVTEHLNLSL-HSLCEVL--GID-----FKQTVVEVHPSLEAEQSK	219
DB	180 ER-KQELKLLQEOQELCEILCMPHYDIDSASVPSLEELNQFQHV----TTLRETKASR	234
QY	220 -----NLSNT-----TIERLAAAANLREMKIORM	244
DB	235 REEFVSIVKRIILCWEELDHTPDTSPERDVVCEDEDAFCLSLNENIATLQKLLRQLEWQKS	294
QY	245 QK---LQDFASSMLELWNLMDTPLEEQOMFQNTICNIAASQEITEPTNTLSTDFLNYYES	301
DB	295 QNEAVCEGLRTQIRELMDRLQIPSEEREAVATI--MSGSKAKVRKA-----LQL	341
QY	302 EVLRLEQLKASKMDLVLKKAEL-----EHRRAHIVGEGYAEFSEIAIEA	351
DB	342 EVDRLEELKQNMKKVIEAIRVELVQWDQCIFYSQEQRAFAPFAEDYTES-----	393
QY	352 GAIDPSIVLEQIEAHIAIVKEEAFSRKIDLEKVERWQACEEAW-----LEDYKDDNR	406
DB	394 -----LLQHDAAEIVRLKNYYEVHKELFEGVQKW-----EETWRLFLEPERKASDNR	441
QY	407 YNAGRGHLLTKRAEKARTLVNKIPGMVDVLRITKIAAWKNERGKEDFTYDGVSLSSMLDE	466
DB	442 F-TNRGGNL-LKEEKQRAKLQKMLPKLEBELKARIELWEQSHSKA-FWNGQKFW-----E	494
QY	467 YMFVROEKEQ-EKKRQDQDKLQDLKAEQALYGSKP-----SPSKPLSTKKAPR	516
DB	495 YVASQWEMHLEKERAQERQKQKQTEMLYGSAPRTSPSKRRGLAPNTPGKARKLNT	554
QY	517 HSMGKANRLSL-----GGATWQPPKTDILHKSVRAAKTEETIGTLPSSSRGLDIAGLP	572
DB	555 TTMSNATANGSIRPIFGTVVHSPVSRIPPSGSKPVAASTCS-GKKTPTGCR-----HGAN	609
QY	573 IKKLSFNASTLRETETPRKPPAQITPGNSVST	605
DB	610 KENLELNGSILLGGYPGSAPLQRFNSINSVAST	642
XX	ADP04701 standard; protein; 639 AA.	
AC	ADP04701;	
XX		
DT	29-JUL-2004 (first entry)	
XX		
DE	Sea squirt protein with tissue specific expression in development Seq296.	
XX	sea squirt; regeneration medicine; gene therapy; cell proliferation;	
KW	differentiation; reproduction; environmental measurement; water survey.	
OS	Ciona intestinalis.	
XX		
PN	JP2004057129-A.	
XX		
PD	26-FEB-2004.	
XX		
PF	31-JUL-2002; 2002JP-00222593.	
XX		

QY 491 LKAEQALYGS---KPSKPLSTTKKAPRHSNGANRRRLSLGGATWQ---PPKTDILHSK 544
Db 1696 LAAERARKQADLEKEELAEELASSLGRNALQDEKRLAEARIAQLEEELEEEQGMAM 1755
QY 545 SVRAAKTTEIEGTLSPSSSRGLDIAGLPKILSFNASTLRETETPRKPPAQITPGNSVSS 604
Db 1756 SDRVRKATQQAELS-----NELATERSTAQKNESARQ---QLERQNKELR 1798
QY 605 TPVRPITNNTEDDENRTPKFTALNPKTPMTVTAPQMAMTFLANKVSATPVSILVYDK- 663
Db 1799 SKLHEMAGV---KSKFKSTIAALEAK-----IALEEQVEQAEAREKQAATKSLKQDKK 1850
QY 664 -PEVTLQEDIDYSFEERRLA 682
Db 1851 LKEILLQVE-----DERKMA 1865

RESULT 15

ABU10398
ID ABU10398 standard; protein; 1945 AA.
XX
AC ABU10398;
DT 20-AUG-2003 (first entry)
XX Human smooth muscle myosin heavy chain (hSMMyHC) variant #1.
DE Human; smooth muscle myosin heavy chain; hSMMyHC; inotropic;
KW antasthmatic; hypertensive; hypotensive; uterine; antiinflammatory;
KW gynaecological; tocolytic; cardiac; vascular disorder;
KW pulmonary disorder; reproductive disorder; immunological disorder;
KW asthma; hypotension; hypertension; urinary incontinence;
KW irritable bowel syndrome; menstrual cramp; premature labor; glaucoma;
KW bronchioconstriction; cardiac malfunction.
XX
OS Homo sapiens.
XX
XX US2003032018-A1.
XX
XX 13-FEB-2003.
XX
XX 10-AUG-2001; 2001US-00927597.
XX
XX 10-AUG-2001; 2001US-00927597.
XX
XX (CYTO-) CYTOKINETICS INC.
XX
XX Malik F, Beraud C, Freedman R, Craven A, Sakowicz R, Hartman J;
XX WPI; 2003-479585/45.
XX N-PSDB; ACA62125.
XX
XX New human smooth muscle myosin heavy chain (hSMMyHC) polypeptide and
PT nucleic acid, useful for diagnosing, preventing or treating e.g. asthma,
PT hypertension, incontinence, menstrual cramps, premature labor or cardiac
PT malfunction.
XX
PS Claim 10; Fig 2A-B; 80pp; English.

XX The invention describes an isolated human smooth muscle myosin heavy
XX chain (hSMMyHC) polypeptide. The hSMMyHC protein, nucleic acid, or its
XX modulator, is useful for diagnosing, preventing or treating vascular,
XX pulmonary, reproductive or immunological disorders. In particular, the
XX hSMMyHC polypeptide, nucleic acid, or its modulator is useful for
XX diagnosing, preventing or treating e.g. asthma, hypotension,
XX hypertension, urinary incontinence, irritable bowel syndrome, menstrual
XX cramps, premature labor, glaucoma, bronchioconstriction, cardiac
XX malfunction, or other medical conditions related to hSMMyHC function. The
XX hSMMyHC polypeptide or nucleic acid is also useful for screening
XX therapeutic agents or hSMMyHC modulators, which may be used for treating
XX the above-mentioned diseases or disorders. This is the amino acid
XX sequence of a human smooth muscle myosin heavy chain (hSMMyHC) variant

SQ Sequence 1945 AA;
Query Match 6.9%; Score 240; DB 6; Length 1945;
Best Local Similarity 20.8%; Pred. No. 7.2e-08;
Matches 154; Conservative 145; Mismatches 313; Indels 128; Gaps 31;
QY 22 ELNVIWVEGPDITRDRMLLELSQ---ECLEVRRKVDQANRSRAQL---RKAIAEGEA 75
Db 1182 EVTVLKALDETRSHQAQVQMRQKHAQAEELTEQLQEPKRAKANLDKKNQTLKXNA 1241
QY 76 ELAGTCSAMGPPVHVHRSNOKLHGLREELNAIPVLEEMKKKKVERWQFVHVE-OIK 134
Db 1242 DLAGELRVIGAKOEVEHKKKLEAQVQELS---KCSGERARAE-LNDKVHKLQNEVE 1297
QY 135 KISSIRPADFVPFVKPVVDQSDLS--LRKLDELTKD-----LESLOKES---D 178
Db 1298 SVTGMINEAGCAIKLAKDVASLSQLQDTQELLQEETROKLVNSTKLQLEERNLQD 1357
QY 179 RLKQVIEHLNLSLCEVIGI-----DFKQTVYEVHPSLDEAEG--SKNLSNNT- 225
Db 1358 QLDEEMAKQNLERHISTLNQLSDSKKQLQDFASTV---EALLEGKRFOKEIENLTQ 1413
QY 226 -IERLAAANRLREMKIQRMQKLODPASSMELNMLMDTPLEEQOMF-----QNITCN 277
Db 1414 QYEEKAAAYDKLETKNRLQQLDLDLVVDLQNRQVLNLEKKQKRFQDLAEKXNISK 1473
QY 278 IA---ASEQITEPTNTLSTDFLNYVESEVLRLEQL-KASRMKDLVLKKALELEHRRRA 332
Db 1474 YADERDRAEAREKETKALSARALEALEAKEELERTNKM-----LKAEMEDLVSSK 1527
QY 333 HLVGEGVAEBSFTEAIEAGIDPSLVLEQIEAHATVKEEAFSKDILEKVERWQNAE 392
Db 1528 DVGKGVHLEKSKRALETOMEEMKTOLEEELEDELOAT-EDAKLRLEV--NMQALKGOPE 1584
QY 393 BEAWLEDYKODNRVYACGAGH---LTLKRAEKARTLV---NKTPGMVDVLTRTIAAWK 445
Db 1585 RDLQARDEQNEKRRQLQRLQHEYTELEDERKQALAAAKKLEGG--DLKDLELQDS 1642
QY 446 NERGKE-----DFTYDGVSLSSMLDEYMFVQKEQE-KKQRDQKQLQDQ 490
Db 1643 AIKGRREAIKQLRKLQAKMDQFQRELEDAERASRDEIFATAKENEKAKSLEADLMQLQSD 1702
QY 491 LKAEQALYGS---KPSKPLSTTKKAPRHSNGANRRRLSLGGATWQ---PPKTDILHSK 544
Db 1703 LAAERARKQADLEKEELAEELASSLGRNALQDEKRLAEARIAQLEEELEEEQGMAM 1762
QY 545 SVRAAKTTEIEGTLSPSSSRGLDIAGLPKILSFNASTLRETETPRKPPAQITPGNSVSS 604
Db 1763 SDRVRKATQQAELS-----NELATERSTAQKNESARQ---QLERQNKELR 1805
QY 605 TPVRPITNNTEDDENRTPKFTALNPKTPMTVTAPQMAMTFLANKVSATPVSILVYDK- 663
Db 1806 SKLHEMAGV---KSKFKSTIAALEAK-----IALEEQVEQAEAREKQAATKSLKQDKK 1857
QY 664 -PEVTLQEDIDYSFEERRLA 682
Db 1858 LKEILLQVE-----DERKMA 1872

Search completed: November 20, 2004, 00:39:17
Job time : 94 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 2004, 00:17:35 ; Search time 26 Seconds
(without alignments)
1759.978 Million cell updates/sec

Title: US-10-619-685-2
Perfect score: 3496
Sequence: 1 MSSAVKDLQHQNSTTCDSEL.....DIDYFERRAIYILARQMV 690

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A.COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B.COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A.COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B.COMB.pep.*
5: /cgn2_6/prodata/1/iaa/6C.COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	240	6.9	1857	4 US-09-917-254-91	Sequence 91, Appl
2	240	6.9	1972	4 US-09-538-092-1084	Sequence 1084, Ap
3	229	6.6	885	2 US-08-533-306A-4	Sequence 4, Appl
4	229	6.6	885	2 US-08-742-923A-4	Sequence 4, Appl
5	228.5	6.5	1972	4 US-08-875-435B-4	Sequence 4, Appl
6	221	6.3	1972	4 US-08-875-435B-3	Sequence 3, Appl
7	219.5	6.3	3210	4 US-09-538-092-1154	Sequence 1154, Ap
8	215.5	6.2	3248	1 US-08-353-700-1	Sequence 1, Appl
9	215.5	6.2	3248	5 PCT-US95-16216-1	Sequence 1, Appl
10	214.5	6.1	2482	1 US-08-328-254-6	Sequence 6, Appl
11	210	6.0	1939	4 US-09-538-092-915	Sequence 915, App
12	209	6.0	1937	4 US-09-538-092-918	Sequence 918, App
13	207	5.9	816	2 US-08-533-306A-6	Sequence 6, Appl
14	207	5.9	816	2 US-08-742-923A-6	Sequence 6, Appl
15	207	5.9	2101	1 US-08-466-390-4	Sequence 4, Appl
16	207	5.9	2101	1 US-08-470-950-4	Sequence 4, Appl
17	207	5.9	2101	1 US-08-467-781-4	Sequence 4, Appl
18	207	5.9	2101	1 US-08-195-487-4	Sequence 4, Appl
19	207	5.9	2101	2 US-08-483-924-4	Sequence 4, Appl
20	207	5.9	2101	3 US-08-452-294-1	Sequence 1, Appl
21	207	5.9	2101	5 PCT-US93-06160-4	Sequence 4, Appl
22	206.5	5.9	3878	4 US-09-914-259-11	Sequence 11, Appl
23	206	5.9	630	4 US-09-248-796A-20275	Sequence 20275, A
24	206	5.9	2349	4 US-09-538-092-914	Sequence 914, App
25	203.5	5.8	2662	4 US-09-595-684B-31	Sequence 31, Appl
26	203.5	5.8	2663	4 US-09-538-092-1252	Sequence 1252, Ap
27	199	5.7	976	4 US-09-538-092-1339	Sequence 1339, Ap

RESULT 1
US-09-917-254-91
; Sequence 91, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 91
; LENGTH: 1857
; TYPE: PRT
; ORGANISM: Homo Sapiens
; US-09-917-254-91

Query Match 6.9%; Score 240; DB 4; Length 1857;
Best Local Similarity 20.8%; Pred. No. 1.2e-09;
Matches 154; Conservative 145; Mismatches 313; Indels 128; Gaps 31;

QY	22	ELNVIWDEVGEPTTTRDRMLLEQ---ECLEVYRKVDQANRSRAQL---RKAIAEGEA	75
DB	1060	EVTVLKALDEBTRSGHEAQVQEMRQKHAQAVEELTEQLQFKRAKANLDKNKQTLKENA	1119
QY	76	ELAGTCSANGPEPPVHVQRQSKLHGLREELNAIVPYLEEMKKKKVRRNOFVHVIE-OIK	134
DB	1120	DLAGELRVLGQAQVEHKKKLEAQVQLQS---KCSGGERARAB-LNDKVHKLQNEVE	1175
QY	135	KISSEIRPADFPVPKVPVQSDLS--LRKLDLDTKD-----LESLOKEKS---D	178
DB	1176	SVTGMLENEAGKAIKADVASLSOLOQTQELLQETQKLVNVTKLQLEERNSLOD	1235
QY	179	RLKQVIEHLNSLHSLCEVLGI-----DFKQTVYVHPSLDABEG--SKNLSNTT-	225
DB	1236	QLDEEMEAQNLERHISTLNIQLSDSKKLQDFASTV---EALSGKKRFOKEIENLTQ	1291
QY	226	-IERLAAAANRLREMKIQRMKLODFASSMLELWNLMDTPLEEQMF-----QNITCN	277
DB	1292	OYEKAAAYDKLEKTKNRLQQLDLDLVLDNQRQLVSNLEKQKQKPDQLLAEKKNISK	1351
QY	278	IA-----ASBOEITEPNTLTDFLNVYSESVLRLEQL-KASKMKDLVLKKKALEEHRRA	332
DB	1352	YADERDRABAREKETKALSARALEEALAEKEELERNKM-----LKAEMEDLVSSK	1405
QY	333	HLVGBEGYAEERFSIEAIEAGIDPSLVLEQIEAHIAITVKEEAFSRKDILEKVERMONACE	392

ALIGNMENTS

Db 1406 DDVGVNHELEKSKRALETKQEMKQLESELELQAT-EDAKLRLEV--NMQALKGQFE 1462
Qy 393 EEAWLEDYKDDNRYNAGRAH---LTLKRAEKARTLV---NKIPGMVDVLRTKIAAWK 445
Db 1463 RDLOARDEQNEEKRRQLRQLHEYTELEDERKQALAAAKKLEG--DLKDLLELQADS 1520
Qy 446 NERKGE-----DFTYDGVSLSSMLDEYMFVRQEKQE-KKRQDQKQLQDQ 490
Db 1521 AIKGREAIKQLKLOAKMDQFQRELEDAKRSDEIFATAKENEKKALEADLMQLOED 1580
Qy 491 LKAEQALYGS---KPSKPLSTKAPRHSMGANRRLSLGGATMQ---PPKTDILHSHK 544
Db 1581 LAAAEARAKQADLEKEELAEELASSLSGRNALQDEKRLLEARIQLEBELEEEQGNMEAM 1640
Qy 545 SVRAAKTTEIGTILSPSSRGLDIAGLPKILSFNASTLTETETPRKPPAQITPGNSVSS 604
Db 1641 SDRVRKATQQAELS-----NELATERSTAQKNESARQ---QLERQNKELR 1683
Qy 605 TPVRPITNTNEDDENRTPKTFTALNPKTPMTVTAPMQMAMTFLANKVVSATPVSIVYDK- 663
Db 1684 SKLHEMAGV---KSKFKSTIAALEAK-----IAQLEEQVEQAREKQATKSLKQDKK 1735
Qy 664 -PEVTLOEDIDYSPERRLA 682
Db 1736 LKEILLQVE-----DERKMA 1750

RESULT 2
US-09-538-092-1084
; Sequence 1084, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormat Version 0.9
; SEQ ID NO 1084
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)-(0)
; OTHER INFORMATION: Polypeptide Accession Number P35749
US-09-538-092-1084

Query Match 6.9%; Score 240; DB 4; Length 1972;
Best Local Similarity 20.8%; Pred. No. 1.3e-09;
Matches 154; Conservative 145; Mismatches 313; Indels 128; Gaps 31;
Qy 22 ELNVIWDEGEPTTTRDRMLLEQ---ECLEVYRKVDQNRSAQL---RKAIAEGEA 75
Db 1175 EYTVLKKALDEETRSHEAQVEMRQKHAQVBELEQEPKRAKANLDKNTLEKENA 1234
Qy 76 ELAGICSAMGEPVHVQRQSNKHLGRLBELNAIVLYSEMKKKVKVERWQFHVIE-QIK 134
Db 1235 DIAGELRVLGAKAQVEVHKKKLELAQVQELQS---KCSGDERARAB-LNDKVHKLQNEV 1290
Qy 135 KISSIRPADFFVFPVVDQSDLS--LRKLDLTKD-----LESQKEKS---D 178
Db 1291 SVTGMLEAEGKAIAKLADVASLSQLQDTQELLQETTRQKLNVTSTKLRLQEEENSQD 1350
Qy 179 RLKQVIEHNSLSLCEVLGI-----DFKOTVVEVHPSLDEAEG--SKNLSNVT- 225

Db 1351 QLDEEMEAQNLERHISTNLNLQSDSKKQLQDPASTV-----EALLEGKKRFQKEIENLTQ 1406
Qy 226 -IERLAAANRLREMKQRMQKLODPASSMLELNLMDTPLEBQOMF-----ONITCN 277
Db 1407 QYEEKAAAYDKLETKQNRLOQELDDLVDNQQLVSNLEKQKQKFDQLLAEEKNTSSK 1466
Qy 278 IA-----ASEQIETPNTLTSTDFLNVSEVSLRLSQL-KASKMKOLVLKKKALEBEHRRRA 332
Db 1467 YADERDRABAEAREKETKALSARALEALFAKEBELRTNKK-----LKAEMEDLVSSK 1520
Qy 333 HLVGEEGYAEFEISAEIAGAIDPSLVLEQIEAHIAIATVKEEAFSRKDILEKVERWQNAE 392
Db 1521 DDVGKNVHELEKSKRALETKQEMKQLESELELQAT-EDAKLRLEV--NMQALKGQFE 1577
Qy 393 EEAWLEDYKDDNRYNAGRAH---LTLKRAEKARTLV---NKIPGMVDVLRTKIAAWK 445
Db 1578 RDLOARDEQNEEKRRQLRQLHEYTELEDERKQALAAAKKLEG--DLKDLLELQADS 1635
Qy 446 NERKGE-----DFTYDGVSLSSMLDEYMFVRQEKQE-KKRQDQKQLQDQ 490
Db 1636 AIKGREAIKQLKLOAKMDQFQRELEDAKRSDEIFATAKENEKKALEADLMQLOED 1695
Qy 491 LKAEQALYGS---KPSKPLSTKAPRHSMGANRRLSLGGATMQ---PPKTDILHSHK 544
Db 1696 LAAAEARAKQADLEKEELAEELASSLSGRNALQDEKRLLEARIQLEBELEEEQGNMEAM 1755
Qy 545 SVRAAKTTEIGTILSPSSRGLDIAGLPKILSFNASTLTETETPRKPPAQITPGNSVSS 604
Db 1756 SDRVRKATQQAELS-----NELATERSTAQKNESARQ---QLERQNKELR 1798
Qy 605 TPVRPITNTNEDDENRTPKTFTALNPKTPMTVTAPMQMAMTFLANKVVSATPVSIVYDK- 663
Db 1799 SKLHEMAGV---KSKFKSTIAALEAK-----IAQLEEQVEQAREKQATKSLKQDKK 1850
Qy 664 -PEVTLOEDIDYSPERRLA 682
Db 1851 LKEILLQVE-----DERKMA 1865

RESULT 3
US-08-533-306A-4
; Sequence 4, Application US/08533306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533,306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869COB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600

TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-533-306A-4

Query Match 6.6%; Score 229; DB 2; Length 885;
Best Local Similarity 21.2%; Pred. No. 2.9e-09;
Matches 160; Conservative 140; Mismatches 311; Indels 142; Gaps 32;

QY 25 VIV-----DEVG--EPDTRDRMLLEQECLEVVYRKV-----DOANR-----SR 63
DB 107 VIVKGMIDQLRDMGCLFEDBERAQOEDALAAQAFEEARRRTREFDRSHREMEAK 166
QY 64 AOL---RKAIAEGEAEAGICSMGEPVHVHVSQKQHLGRLREELNAIVPYLEEMKKVK 120
DB 167 ANLDKKNQTLKENADLAGELRVLGQAQVEHKKKKLEAQVQELQS---KCSGGERARA 223
QY 121 ERNQVFHVIE--QIKKISSEIRPADVPKVPVVDQSDLS--LRKLDLTKD-----168
DB 224 E-LNDKVHKLQNEVESVTGMLNEAEGKAIKLAQDVASLSQDQDTQELLQEEETROKLNVS 282
QY 169 --LESQKEKS---DLKQVIEHLNLSLSLCEVLGI-----DFKQTVVYVHPSLD 213
DB 283 TKLRQLEEBRNLSQDQDEMEAKQNLERHISTLNQLSDSKKKLQDFASTV---EAL 338
QY 214 EAG--SKNLSNTT--IERLAAAANRLREMKIORMQKQDFASSMELNMLMDTPLEEQ 269
DB 339 EGKRFQKEIENLTQYEEKAAAYDKLETKNRLQELDDLVDLDNQRLVSNLEKKQR 398
QY 270 MF-----QNTICNTA---ASEQITEPTNTLSTDFLNVSEVLRLEQL--KASQKDL 317
DB 399 KFDQLLAEEKNISSKVADERDRAEAAREKETKALSARALEBALEKEELERTNKN---455
QY 318 VLKKALEHRRRAHLVGEEGAEEFSIEAEGAIQPSLVLEQIEAHATVKEBAFSR 377
DB 456 ---LKAEMEDLVSSKDDVGKQNVHELEKSKALETQMEEMKTQLELEDEL--QASEDAKLR 511
QY 378 KDILEKVERWQNAACEEAWLEDYKDDNRYNAGRGH-----ITLKRAEKARTLVNKP 432
DB 512 LEV--NQALKQGFERDLQARDQNEKRLQRLQHEYTELEDEDERNERALAAAKKL 569
QY 433 MVDVLRTKIAAKNRGEK-----DFTYDGVSLSSMLDEYMFVYRQKEQE--477
DB 570 EGBLKDLQADSAIKGRBAIKQLRKLQAKMDFQRELEDAASRDEIFATAKENEKKA 629
QY 478 KKRQDKKLODQKAEQALYGS---KPSFKPLSTKKAPRHSMGANRLSLGGATMQ 534
DB 630 KSLEADLMQLEDLAAARARKQADLEKEBELASSUSGRNALQDEKRLREARTIAQLE 689
QY 535 ---PPKTDILHKSVAAKTEETIGTLPSSSRGLDIAGLPIKKLSFNASTLRETETPRK 591
DB 690 EELEEGCNMEASDRVKATQAEQLS-----NELATERSTAQKNESARQ 735
QY 592 PPAQITPGNSVSTPVPIPTNTEDDENRTPKTTFTALNPKTPMTVTAPMQMAMTPSLANK 651
DB 736 ---QLERQNKRLSKLHEMEGAV---KSKFKSTIAALEAK-----IAQLEEQVEQAREK 784
QY 652 VSATPSLVYDK--PEVTIQEDIDYSEERLA 682
DB 785 QAATKSUKQDKKLKEILOVE-----DERKMA 812

RESULT 4

US-08-742-923A-4
; Sequence 4, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.

APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/742,923A
FILING DATE: No. 5869611 member 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Dean F.
REGISTRATION NUMBER: 36583
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 885 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923A-4

Query Match 6.6%; Score 229; DB 2; Length 885;
Best Local Similarity 21.2%; Pred. No. 2.9e-09;
Matches 160; Conservative 140; Mismatches 311; Indels 142; Gaps 32;

QY 25 VIV-----DEVG--EPDTRDRMLLEQECLEVVYRKV-----DOANR-----SR 63
DB 107 VIVKGMIDQLRDMGCLFEDBERAQOEDALAAQAFEEARRRTREFDRSHREMEAK 166
QY 64 AOL---RKAIAEGEAEAGICSMGEPVHVHVSQKQHLGRLREELNAIVPYLEEMKKVK 120
DB 167 ANLDKKNQTLKENADLAGELRVLGQAQVEHKKKKLEAQVQELQS---KCSGGERARA 223
QY 121 ERNQVFHVIE--QIKKISSEIRPADVPKVPVVDQSDLS--LRKLDLTKD-----168
DB 224 E-LNDKVHKLQNEVESVTGMLNEAEGKAIKLAQDVASLSQDQDTQELLQEEETROKLNVS 282
QY 169 --LESQKEKS---DLKQVIEHLNLSLSLCEVLGI-----DFKQTVVYVHPSLD 213
DB 283 TKLRQLEEBRNLSQDQDEMEAKQNLERHISTLNQLSDSKKKLQDFASTV---EAL 338
QY 214 EAG--SKNLSNTT--IERLAAAANRLREMKIORMQKQDFASSMELNMLMDTPLEEQ 269
DB 339 EGKRFQKEIENLTQYEEKAAAYDKLETKNRLQELDDLVDLDNQRLVSNLEKKQR 398
QY 270 MF-----QNTICNTA---ASEQITEPTNTLSTDFLNVSEVLRLEQL--KASQKDL 317
DB 399 KFDQLLAEEKNISSKVADERDRAEAAREKETKALSARALEBALEKEELERTNKN---455
QY 318 VLKKALEHRRRAHLVGEEGAEEFSIEAEGAIQPSLVLEQIEAHATVKEBAFSR 377
DB 456 ---LKAEMEDLVSSKDDVGKQNVHELEKSKALETQMEEMKTQLELEDEL--QASEDAKLR 511
QY 378 KDILEKVERWQNAACEEAWLEDYKDDNRYNAGRGH-----ITLKRAEKARTLVNKP 432
DB 512 LEV--NQALKQGFERDLQARDQNEKRLQRLQHEYTELEDEDERNERALAAAKKL 569
QY 433 MVDVLRTKIAAKNRGEK-----DFTYDGVSLSSMLDEYMFVYRQKEQE--477

Db 570 EGDLDLEQADSAIKGEEAIKQLKLAQMKDFORELEADARASGRDEIFATAKENEKA 629
Qy 478 KKRQDOKLQDLKAEQALYGS---KPSPKPLSTKAPRHSWGCGANRLSLGATWQ 534
Db 630 KSLADLMQLQDLAAERARKQADLEKEELAEELASSLSGRNALQDEKRLREARLAQLE 689
Qy 535 ---PPKTDILHKSVAARAKTTEIGTLPSSSRGLDIAGLPKLISFNASTLRETETPRK 591
Db 690 ELEBEQGNWEAMSDVRKATQAEQLS-----NEIATERSTAKNESARQ 735
Qy 592 PFAQITPGNSVSTPVRPITNNEDDENRTPKTFTALNPKTMTVTAPMOMTSPSLANK 651
Db 736 ---QLERQNKELRSKLHEMEGAV---KSKFKSTIAALEAK-----IAQLEEQVEQEAEREK 784
Qy 652 VSATPVSIVYDK--PEVTLOEDIDYSPERRLA 682
Db 785 QAATKSLKQDKKLKEILLQVE-----DERKWA 812

RESULT 5

US-08-875-435B-4
; Sequence 4, Application US/08875435B
; Patent No. 6593304
; GENERAL INFORMATION:
; APPLICANT: Hasegawa, Kazuhide
; APPLICANT: Arakawa, Emi
; APPLICANT: Oda, Shoji
; APPLICANT: Matsuda, Yuzuru
; APPLICANT: Takahashi, Katsuhito
; APPLICANT: Sugahara, Michihiro
; APPLICANT: Ishiyama, Haruo
; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING
; TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO
; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND
; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE
; TITLE OF INVENTION: RECOMBINANT DNA
; FILE REFERENCE: 07898-013001
; CURRENT APPLICATION NUMBER: US/08/875,435B
; PRIOR FILING DATE: 1997-07-25
; PRIOR APPLICATION NUMBER: PCT/JP96/00134
; PRIOR FILING DATE: 1996-01-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-08-875-435B-4

Query Match 6.5%; Score 228.5; DB 4; Length 1972;
Best Local Similarity 19.6%; Pred. No. 1e-08;
Matches 151; Conservative 159; Mismatches 316; Indels 143; Gaps 30;
Qy 4 AVKDQLHQ-MSTCDLSLL-----ELNVINWDEVEGPDTRDRMLLELEQ---ECLEYVR 53
Db 1150 ALKTELEDLTDTATQOELRAKQCEVTVLKKALDEETRSHEAQVQEMRQKHTQVVELT 1209
Qy 54 KVDQANRSRAQL---RKAIAEGEAEIAGICSAMGCEPPVHYR----- 92
Db 1210 EOLEQFKRAKANLDTKTQLEKENADLAGELRVLQAKQVEHKKKLEVLQELQSKCS 1269
Qy 93 -----OSNQKLHGLREBELNAIVPYLEMKKKKVERWNQFVHVTEIQIKKTSSEIRPADF 145
Db 1270 DGERARAEINDVKHKLQNEVESVTGMLSEAEKAIK-----LAKEVASIGSLODQOE 1322
Qy 146 VPFKVPVQSDLSLRLKDELTKDLSLOKESDRL--KQVIE-HUNSLHSICEVLGIDFK 202
Db 1323 LIQETETQKLVNS-TKLQLEDERNSLOQLDEEMAEKQNLERHISTLN-----IQLSDSK 1377
Qy 203 QTVYEVHSLDEAGSKNSLNTTIERL-----AAAANRLREMKIQRMQKLOQFASML 255
Db 1378 KKLQDFASTVESLESGKKRFQKEIBSLTQQVEEKAAAYDKLEKTKNRLQQLDLDLVLDL 1437

Qy 256 ELWNLMDTPEEQMF-----QNITCNIA-----ASQEIETPTNTLSTDFLNVYSESVL 304
Db 1438 NORQLVSNLEKQKQKFDQLAEEKNISSKYADEDRDAEAEAREKETKALSARALEEAL 1497
Qy 305 RLEQL-KASKMKOLVLKKAELHEHRRRAHLVBEGBVAEFSIFAIEAGADPSLVLEQI 363
Db 1498 AKEELERTNKM-----LKAEMEDLVSSKDDVGNVHELEKSKRALETQMEEMKTQLEEL 1551
Qy 364 BAHITATVKEEAFSRKDILEKVERWQACEEAEWLDEYNKDDNRYNAGRGAAH---LTLKRA 420
Db 1552 EDELQAT-EDAKLRELV--NMQALKVQFERDQARDEQNEERKQRLQQLHEYTELEDE 1608
Qy 421 EKARTLV---NKIPGMVDVLTAKIAWKNERGKEDFTYDGVSVLSMDLDEYMFVROE--- 473
Db 1609 RKQRALAAAKKLEG--DLKDLELDQADSAIKGEEAIKQLKLAQMKDFORELEADARA 1666
Qy 474 -----KEQEKCR---QRDQKQLQDLKAEQALYGS---KPSPKPLSTKAPRHS 518
Db 1667 SRDEIFATAKENEKAKSLBADLMQLQEDLAAERARKQADLEKEELAEELASSLSGRNA 1726
Qy 519 MGGANRLSLGGATMQ---PKTDILHKSVAARAKTTEETIGTLPSSSRGLDIAGLPKK 575
Db 1727 LQDEKRLREARIAQLEEBEEOGNMEAMSDVRKATQAEQLS-----NE 1772
Qy 576 LSFNASTLRETETPRKPFQAQITPGNSVSTPVRPITNNEDDENRTPKTFTALNPKTDMT 635
Db 1773 LATERSTAKNESARQ---QLERQNKELSKLQEMEGAV---KSKFKSTIAALEAK----- 1822
Qy 636 VTAPMOMTSPSLANKVSATPVSIVYDK--PEVTLOEDIDYSPERRLA 682
Db 1823 -IAQLEEQVEQEAEREKQAAAKALKQDKLKEMLLQVE-----DERKWA 1865

RESULT 6

US-08-875-435B-3
; Sequence 3, Application US/08875435B
; Patent No. 6593304
; GENERAL INFORMATION:
; APPLICANT: Hasegawa, Kazuhide
; APPLICANT: Arakawa, Emi
; APPLICANT: Oda, Shoji
; APPLICANT: Matsuda, Yuzuru
; APPLICANT: Takahashi, Katsuhito
; APPLICANT: Sugahara, Michihiro
; APPLICANT: Ishiyama, Haruo
; TITLE OF INVENTION: RECOMBINANT DNA COMPRISING DNA CODING
; TITLE OF INVENTION: FOR MYOSIN HEAVY CHAIN SM1 ISOFORM PROTEIN INSERTED INTO
; TITLE OF INVENTION: VECTOR DNA, MICROORGANISM CARRYING THE RECOMBINANT DNA, AND
; TITLE OF INVENTION: AN AGENT FOR TREATMENT OF ARTERIOSCLEROSIS COMPRISING THE
; TITLE OF INVENTION: RECOMBINANT DNA
; FILE REFERENCE: 07898-013001
; CURRENT APPLICATION NUMBER: US/08/875,435B
; CURRENT FILING DATE: 1997-07-25
; PRIOR APPLICATION NUMBER: PCT/JP96/00134
; PRIOR FILING DATE: 1996-01-25
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-875-435B-3

Query Match 6.3%; Score 221; DB 4; Length 1972;
Best Local Similarity 21.0%; Pred. No. 3.9e-08;
Matches 159; Conservative 133; Mismatches 302; Indels 164; Gaps 32;
Qy 22 ELNVINWDEVEGPDTRDRMLLELEQ---ECLEYVRVRRVQDANRSRAQLRKA---IAEGEA 75
Db 1175 EVTVLKKALDEETRSHEAQVQEMRQKHTQVVELEQLEQFKRAKANLDSKQTLKENA 1234
Qy 76 ELAGICSAMGCEPPVHVHVRQSNKL-----HGLUREELNAIVP 110

1235 DLAGELVLGQAQKEVEHKKKLEVLQDLQSKSDGERARAEKSKVHKLVNEVESVTG 1294
111 YLEEMKKKKVVERWQVHVEIQIKKISSIRPADFPVFKVVDQSDL-----SLRKDELDT 166
1295 MLNBAEGKAIK-----LAKDVASIGSLQ-----DTQELQETROKLVNST 1336
167 KDLSELOKESK-----DLKQVIEHLNSHLSLCEVLGI-----DFKQTVVEHPSLD 213
1337 K-LRQLEDERNLQDQDEMAEKQNLERHVSNTLNQLSDSKKQLQDFASTIEVME----- 1391
214 BAEGSKNLS-----NTTIERLAAANRLREMKIORMQKLODPASSMELMNMMDTPLEE 267
1392 --EGKKRLQKEMBLGSOQVEEAAAYDKLEKTKNRLQQLDLDLVLDNQRQLVSNLEKK 1449
268 QQMF-----QNTICNIA-----ASEQIEITPNTLSTDFLNVYSEVRLQLEQ-KASKMK 315
1450 QKFDQLLAEKKNISSYADERDRAEAREKETKALSARALEEALAELEKELERNKM- 1508
316 DLVLKKAELERHRRRAHLVGBEGYAEFSEIAEAGADPSLVLEQIEAHIAIVKEEAF 375
1509 -----LKAEMEDLVSSKDDVGKQVHELEKSKRALETQMEEMKTQLESEDDVQAT-EDAK 1562
376 SRDKILEKVERWQNAEEAEWLDYNDKDDNRVYNAGRAH-----LTLKRAEKARTLV-----N 428
1563 LRLEV--NQALKGQFERDLQARDEQNEEKRRQLQRLHEYTELEDERKQALAAAAKK 1620
429 KIPGMVDVLRKTAANKNBERGKEDFTYDGVSLSSM-----LDEYMFVROE-----KEQ 476
1621 KLEG--DLKDLQADSAIKGREBAIKQLRKLOAQMKDFQRELDADARASDEIFATSKEN 1678
477 EKKE--QRDQKLODLKAEQALYGS---KSPSPKPLSTKKAPRHSMGANRRLSLGG 530
1679 EKAKSLEADLMQLEDLAAERARKQADLEKEELAEELASSUGRNTLODEKRLERAI 1738
531 ATMPPKPTDILHKSVRAAKTEBEIGTLPSSSRGLDIAGLPKIKLSFNASTLRET----- 586
1739 AQLE-----EELSEEQGNMEAMSDR-VKATLQAEQLSNELATERSTAQKN 1783
587 ETPRKPAQITPGNSVSTPVRITNTNEDDENRPTFTALNPKPTMTVTPMQAMTP 646
1784 ESARQ---QLERQNKELRSKLQSEVAV---KAKLKSTVAALAK-----IAQLEBQVEQ 1832
647 SLANKVSATPSLVLYDK--PEVTLQEDIDYSPFERLIA 682
1833 EAREQAQATSKLKQKOKKLEVLQVE-----DERKWA 1865

RESULT 7

US-09-538-092-1154
; Sequence 1154, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1154
; LENGTH: 3210
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P49454

US-09-538-092-1154

Query Match 6.3%; Score 219.5; DB 4; Length 3210;
Best Local Similarity 20.6%; Pred. No. 1.1e-07;
Matches 139; Conservative 124; Mismatches 256; Indels 155; Gaps 29;
QY 1 MSSAVKQDLH-----QMSTTCDSLLLELNVIVDEVEPEPTTRDRMLLELEOELEVEYRKV 56
DB 2119 VSELLKDKTHLQELQKLSLEKDSQALSLTKCELENOIAQNLKEKELLVXSESLSQA---RL 2175
QY 57 DOANRSRAQLKATIAEGEAGELAGICSAMGEPVHVVRQSNOKLHGRLEELNATVPVLEEMK 116
DB 2176 SESDYKLVNLSKAL---EALV-----EKGEFALRLSSQTQEVVQLRGIEKLRVLEADE 2228
QY 117 KKKVERMQFVHVIRIQQK-----ISSEIRPADPV 146
DB 2229 KQK-----LHIAEKLERENDSLKDKVENLERELQMSBENQELVILDAENSCABVE 2281
QY 147 PKFVPVDSLSRLKDLDTKDLSELOKESDRKQVIEHLNSHLSLCEVLGIDFKQTVY 206
DB 2282 TLTKTQIEEMARSL-KVFEL--DLVTLRSEKENLTKQIQEQQLSSELDKLLS-SFKSLLE 2337
QY 207 EVHPSLDAE-GSKNLSNTTIERLAAANRLREMKIORMQKLODPASSMELMNMMDTFL 265
DB 2338 E-----KEQABIQIKESKTAVEML---QNQLKELN-EAVALCGDQEIEMKATEQSLDPPI 2389
QY 266 BEQQMFON-----ITCNIAASEQIEITPNTLSTDFLNVYSEVRLLEQLKASK-----MKD 316
DB 2390 ESEHQLRNSIEKLRARLEADEKK-----QLCVLQQLKESSEHHDLLKG 2432
QY 317 LVLKKAELERHRRRAHLVGBEGYAEFSEIAEAGADPSLVLEQIEAHIAIVKEEAFS 376
DB 2433 RVNLERELEIARTQNEHAALAEANSKGEVETLKAKIEGMTQSLRGLELDVVVTIRSE--- 2489
QY 377 KDIILEKVERWQNAEEAEWLDYNDK-DNRVYNAGRAHLLTKRAEKARTLVNKPIMGWD 435
DB 2490 KEDLTNELQEQERISE---LEIINSSPENILQEQEKQVQMK---EKSTANEMLTQQLK 2544
QY 436 VLRTKIAAWKN-----RGKED-----FTYDGVSLSSMLDE---YMFVR----- 471
DB 2545 ELNERRVAAHLNDQEAACKAEQNLSSQVECLELEKALQQLGLDEAKNNYIVLQSSVNGLIQ 2604
QY 472 -----QEKEQKQRQDKLQDLKAEQALYGSKPSKPLSTKKAPRHSMGGA 522
DB 2605 EVEDGKQKLEKDEEISRLKNQIQEQELVSKLSQVEG-----EHQL-WK 2648
QY 523 NRRLSLGGATMQ-PPKTDILHKSVRAAKTEBEIGTLPSSSRGLDIAGLPKIKLSF--- 578
DB 2649 EQNLERLNLTVELEQKIQVLQSKNASLQDTLTVLQSSYKNLENELELT--KMDKMSFVEK 2706
QY 579 -NASTLRETETPRK 591
DB 2707 VNMKTAKETELQRE 2720

RESULT 8

US-08-353-700-1
; Sequence 1, Application US/08353700
; Patent No. 5599919
; GENERAL INFORMATION:
; APPLICANT: YEN, TIMOTHY J.
; APPLICANT: RATTNER, JEROME B.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING A
; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,
; TITLE OF INVENTION: AND METHODS OF USE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN
; STREET: 1601 MARKET STREET, SUITE 720
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2307

207 EVHPSLDEAE-GSKNLSNTTIERLAAANRLREMIQRMOKLQDFASSMELNMDTLP 265
2338 E-----KEQAEIQKEESKTAVEML-----QNQLKELN-EAVALCGQOEIMKATEQSILDPPI 2389
266 EQQQMFON-----ITCNIAASEQIEPNTLSTDFLNYVESEVLRLEQLKASK-----MKD 316
2390 EEEHQLRNSIEKLRARLEADEKK-----QLCVLQQLKESEHHADLLKG 2432
317 LVLKKAELIEEHRRAHLVGESEAEFSEIAEAGADPSLVLEQIEAHIAIVKKEAFS 376
2433 RVENLERELEIARTNQEHAALEAENSKEVETLKAKIEGMTQSLRGLELDVVITRSE-----2489
377 RDIILEKVRWQACEEAEAWLEDYKND-DNRYNAGRGHILTKRAEKARTLVNKIPGMVD 435
2490 KENLTNELQKEQERISE---LEIINSFENILQKEQEKVQMK--EKSSTAMEMLQTLQK 2544
436 VLRTKIAAWKNE-----RGKED-----FTYDGVLSLSMLDE-----YMFVR-----471
2545 ELNERVAALHNDQEAACKAEQNLSSQVECLELEKAKQLLQGLDEAKNNYIVLQSSVKGLIQ 2604
472 -----QEKEQEKRRDQKQDOLKAEQAEALYGSKPSKPLSTKKAPRHSMGGA 522
2605 EVEDGKQKLEKDEEISRLKNQIQDQELVSKLSQVEG-----EHQL-WK 2648
523 NRRLSLGGATMO-PPKTDILHSKSVRAAKKTEIEIGTLSPSSRGDLIAGLPKILKLSF---578
2649 EQNLEIRNLTVLEQKIQLVQSKNASLQDTLEVLOQSSYKNLENELELT--KMDKMSFVEK 2706
579 -NASTLRETETPRK 591
2707 VNKMTAKETELQRE 2720

RESULT 10

US-08-328-254-6
Sequence 6, Application US/08328254
Patent No. 5710022
GENERAL INFORMATION:
APPLICANT: Zhu, Xueliang
APPLICANT: Lee, Wen-Hwa
TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/328,254
FILING DATE: 24-Oct-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/141,239
FILING DATE: 22-Oct-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-CJ 1191
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2482 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-328-254-6
Query Match
6.1%; Score 214.5; DB 1; Length 2482;
Best Local Similarity 20.5%; Pred. No. 1.8e-07;
Matches 138; Conservative 125; Mismatches 256; Indels 155; Gaps 29;
1 MSSAVKDDQLH---QMSTTCDLSLLELVNWDVEGPEPTTRDRMLLELEQECLEVYRRKV 56
1391 VSELLKDKTHLQEKLSLEKDSQALSSTKCELENQIAQLNKEKELLVXSESLOA---RL 1447
57 DOANRSRAQLKATAGAEAGELAGICSANGEPVHVHQSNQKLGHLGIEELNAIVPYLEMK 116
1448 SESDYKLVNSKAL---EALV---EKGEFALRSSTQSEVHQURRGIEKURVRIEAD 1500
117 KKKVERWQFVHVIOIKK-----ISSEIRPADFV 146
1501 KKQ-----LHIAEKLRERENDSLKDKVENLERELQMSSEENQELVILDENSCAEVE 1553
147 PFKVPVDSLSRLKDLBTLKDSLOKEKSDRLKQVIEHLNLSHLSCEVLGIDFKQTVY 206
1554 TLKTQIEEMARSL-KVFEL--DLVTLRSEKENLTQIQEKQQLSELDKLLS-SFKSLLE 1609
207 EVHPSLDEAE-GSKNLSNTTIERLAAANRLREMIQRMOKLQDFASSMELNMDTLP 265
1610 E-----KEQAEIQKEESKTAVEML---QNQLKELN-EAVALCGQOEIMKATEQSILDPPI 1661
266 EQQQMFON-----ITCNIAASEQIEPNTLSTDFLNYVESEVLRLEQLKASK-----MKD 316
1662 EEEHQLRNSIEKLRARLEADEKK-----QLCVLQQLKESEHHADLLKG 1704
317 LVLKKAELIEEHRRAHLVGESEAEFSEIAEAGADPSLVLEQIEAHIAIVKKEAFS 376
1705 RVENLERELEIARTNQEHAALEAENSKEVETLKAKIEGMTQSLRGLELDVVITRSE---1761
377 RDIILEKVRWQACEEAEAWLEDYKND-DNRYNAGRGHILTKRAEKARTLVNKIPGMVD 435
1762 KENLTNELQKEQERISE---LEIINSFENILQKEQEKVQMK--EKSSTAMEMLQTLQK 1816
436 VLRTKIAAWKNE-----RGKED-----FTYDGVLSLSMLDE-----YMFVR-----471
1817 ELNERVAALHNDQEAACKAEQNLSSQVECLELEKAKQLLQGLDEAKNNYIVLQSSVNGLIQ 1876
472 -----QEKEQEKRRDQKQDOLKAEQAEALYGSKPSKPLSTKKAPRHSMGGA 522
1877 EVEDGKQKLEKDEEISRLKNQIQDQELVSKLSQVEG-----EHQL-WK 1920
523 NRRLSLGGATMO-PPKTDILHSKSVRAAKKTEIEIGTLSPSSRGDLIAGLPKILKLSF---578
1921 EQNLEIRNLTVLEQKIQLVQSKNASLQDTLEVLOQSSYKNLENELELT--KMDKMSFVEK 1978
579 -NASTLRETETPRK 591
1979 VNKMTAKETELQRE 1992

RESULT 11

US-09-538-092-915
Sequence 915, Application US/09538092
Patent No. 6753314
GENERAL INFORMATION:
APPLICANT: Giot, Loic
APPLICANT: Mansfield, Traci A.
TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
FILE REFERENCE: 15966-542
CURRENT APPLICATION NUMBER: US/09/538,092
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: 60/127,352
PRIOR FILING DATE: 1999-04-01
PRIOR APPLICATION NUMBER: 60/178,965
PRIOR FILING DATE: 2000-02-01
NUMBER OF SEQ ID NOS: 1387
SOFTWARE: CuraPatSeqFormatter Version 0.9


```
Patent No. 5837457
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/533,306A
FILING DATE: September 25, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869COB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-533-306A-6

Query Match 5.9%; Score 207; DB 2; Length 816;
Best Local Similarity 20.6%; Pred. No. 1.3e-07;
Matches 153; Conservative 133; Mismatches 295; Indels 160; Gaps 30;

QY 31 GEPDTRDRMLLEQECLEVYR-----KVD-----QANRSRAQL 66
DB 74 GEORQTPSREYVDLEREAGVKYLKAPMLNGVCVIMQWIDLQRLDGMGCLFDEERAQ 133
QY 67 RKAIAEGEAEAGICSAAGPPVHVRSQSNOKLHGLREEL-----NAIVPYLEEMKKKKVER 122
DB 134 EDALQ-----QAFEBARRTRFEFDRDRSHREEMENEVESVTGMLNEAEGKAIK- 193
QY 123 WNOFVHVIEQIKKISSIRPADEVFPKVPVDQSD-----SLRKLDELTKDLESLOKEKS- 177
DB 184 -----LAKDVASISSQLQ-----DTQELLQEETRQKLVNSTK-LRQLEERN 225
QY 178 --DRLKQVIEHLNLSLCEVLGI-----DFKQTYVEVHPSLDEAEG--SKNLSN 223
DB 226 LQQLDDEMEAKONLHRIHTNLINLSQSKKLQDFASTV-----EALLEGKRFQKEIEN 281
QY 224 TT--IERLAAANRLREMKIQKQKQDFASSMELWNLMWDTPLLEQOMF-----QNI 274
DB 282 LTQOYEEKAAYDKLETKRNLQOELDLVVDLQNRQLVSNLEKKQKQKFDOLLAKEKNI 341
QY 275 TCNIA-----ASEQIEITPNTLSTDFLNYSEVLRLEQL-KASKMKDLVLKKXAELEHR 329
DB 342 SSKYADERDRAEAAREKETKALSARALEEALEAKEELERTNKM-----LKAEMEDLV 395
QY 330 RRAHVGEEGVAEFSFEATEAGNAIDPSLVLEQIEAHIAIVKEAFSRKDILEKVERQW 399
DB 396 SSKDDVGKNVHELEKSKRALETOMEEMKTQLELEDEL-QASEDAKLRLLEV--NMQALG 452
QY 390 ACBEEAWLEDYNDKDDNRYNAGRGH-----LTLKRAEKARTLVNKPFGWVDVLRTKIAAW 444

Db 453 QFERDLQARDEQNEKRRQLQRLHBYETELEDERNERLALAAAKKKLEGDUKLELQAD 512
QY 445 KNERGKE-----DFTYDGVSLSSMLDEYMFVROEKEQ-KKQRQDKKQD 489
Db 513 SAIKGREAEAIQRLKLOQMCKDFQFELEDARASRDEIFATAKENEKAKSLEADLMQLQE 572
QY 490 QLKAEQEAALYGS---KPSPSKPLSTKAPRHSMGANRRSLGGATMQ---PPKTDILHS 543
Db 573 DLAAARARKQADLEKEELABELASSLSGRNALQDEKRLLEARIAQLEEELEEEQGNMEA 632
QY 544 KSVRAAKTEETIGTSLSPSSRGCLDIAGLPKVLSPNASTLTRETETPRKPAQITPNSVVS 603
Db 633 MSDRVKATQQAQELS-----NELATERSTAQKNESARQ---QLERQNKEL 675
QY 604 STPVRPITNNTEDDENRTPKFTALNPKTPMTVTAPMQMAMTPSLANKVYSATPVSIVYDK 663
Db 676 RSKLHEMGA---KSKFKSTIAALEAK-----IAQLEEQVEQEAEREKQATKSLKQDK 727
QY 664 --PEVTLOQEDIDYSFEERRLA 682
Db 728 KLKEILLQVE-----DERKMA 743

RESULT 14
US-08-742-923A-6
Sequence 6, Application US/08742923A
Patent No. 5869611
GENERAL INFORMATION:
APPLICANT: Liu, Pu
APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.
APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16
TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
STREET: P.O. Box 828
CITY: Bloomfield Hills
STATE: MI
COUNTRY: USA
ZIP: 48303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,923A
FILING DATE: No. 5869611ember 1, 1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869DVC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 816 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-923A-6

Query Match 5.9%; Score 207; DB 2; Length 816;
Best Local Similarity 20.6%; Pred. No. 1.3e-07;
Matches 153; Conservative 133; Mismatches 295; Indels 160; Gaps 30;

QY 31 GEPDTRDRMLLEQECLEVYR-----KVD-----QANRSRAQL 66
DB 74 GEORQTPSREYVDLEREAGVKYLKAPMLNGVCVIMQWIDLQRLDGMGCLFDEERAQ 133
QY 67 RKAIAEGEAEAGICSAAGPPVHVRSQSNOKLHGLREEL-----NAIVPYLEEMKKKKVER 122
DB 134 EDALQ-----QAFEBARRTRFEFDRDRSHREEMENEVESVTGMLNEAEGKAIK- 193
QY 123 WNOFVHVIEQIKKISSIRPADEVFPKVPVDQSD-----SLRKLDELTKDLESLOKEKS- 177
DB 184 -----LAKDVASISSQLQ-----DTQELLQEETRQKLVNSTK-LRQLEERN 225
QY 178 --DRLKQVIEHLNLSLCEVLGI-----DFKQTYVEVHPSLDEAEG--SKNLSN 223
DB 226 LQQLDDEMEAKONLHRIHTNLINLSQSKKLQDFASTV-----EALLEGKRFQKEIEN 281
QY 224 TT--IERLAAANRLREMKIQKQKQDFASSMELWNLMWDTPLLEQOMF-----QNI 274
DB 282 LTQOYEEKAAYDKLETKRNLQOELDLVVDLQNRQLVSNLEKKQKQKFDOLLAKEKNI 341
QY 275 TCNIA-----ASEQIEITPNTLSTDFLNYSEVLRLEQL-KASKMKDLVLKKXAELEHR 329
DB 342 SSKYADERDRAEAAREKETKALSARALEEALEAKEELERTNKM-----LKAEMEDLV 395
QY 330 RRAHVGEEGVAEFSFEATEAGNAIDPSLVLEQIEAHIAIVKEAFSRKDILEKVERQW 399
DB 396 SSKDDVGKNVHELEKSKRALETOMEEMKTQLELEDEL-QASEDAKLRLLEV--NMQALG 452
QY 390 ACBEEAWLEDYNDKDDNRYNAGRGH-----LTLKRAEKARTLVNKPFGWVDVLRTKIAAW 444
```

Db 74 GEORQTPSYREYDLEREAGKYLKAPMILNGVCVIWGLDQLDGMCLFDFBERAQ 133
Qy 67 RKAIAEGEAEAGLACISAMGEPVHVHVSQKQLHGLREEL-NAIVPVLKEMKKKVER 122
Db 134 ERLAQ-----QAFBARRRREFDRDRSHREEMENEVSTGMLNEAGKAIK- 183
Qy 123 WQFVHVIBQIKKISSEIRPADVPFVKVVDQSDL-----SLRKLDELTKDLESLOKES- 177
Db 184 -----LAKDVASLSQLQ-----DTQELLQEBTRQKLVNYSK-LRQLEBEENS 225
Qy 178 --DRLKQVIEHLNSLSLCEVLGI-----DPKQTVYEVHPSLDEAG--SKNLSN 223
Db 226 LQDQLEENEAONLHERHISTINLIQSDSKKLQDFASTV-----BALBEGKRKFQKEIEN 281
Qy 224 TT--TERLAAANRLREMIORMOKLQDFASMLLWNLMDTPLEEQMF-----QNI 274
Db 282 LTQVVEKAAVYKLEKTKNRLQQLDLDLVLDNQRLVSNLEKKQRKFDQLLAEEKNI 341
Qy 275 TCNIA---ASQEITEPNTLSTDFLNVYSEVLRLEQL-KASKMKDLVLKKAELBEHR 329
Db 342 SSKYADERDRAEAREKETKALSARALEAEALKEELERTNM-----LKAEMEDIV 395
Qy 330 RAHLVGEYGEYAEBSIEAIEAGADPSLVLEQIEAHATVKEBAFSPKDIIEKVERQW 389
Db 396 SSKDDVGKXVHELEKSKRALETQMBEEMKTQLEDEDEL-QASEDAKRLV--NMQALKG 452
Qy 390 ACEEEAWLDYKNDNRNAGRGAH-----LTLKAEKARTLVNKPIMGVVDVLRTKIAAW 444
Db 453 QFERDQARDEQNEKRRQLQOLHEYTELEDEBERNARALAAAKKLEGDLKDELOAD 512
Qy 445 KNERGKE-----DFTYDGVSLSSMLDEYMFVROEKEQ-KKQRDOKQLQD 489
Db 513 SAIKGREEAIKQLKLAQMKDFORELEDAARSRDEIFATAKENEKAKSLEADLMQLQE 572
Qy 490 QLAKEQEAALYS---KPSPSKPLSTYKAPRSHMGANRLSLGGATMQ---PPKTDILHS 543
Db 573 DLAAERARKQADLEKEALEELASLSGRNALQDEKRLLEARIQALEEELEEEQGNMEA 632
Qy 544 KSVRAAKTEETIGTSPSSSRGLDIAGLPIKKLSFNASTLRETETPRKPAQITPGNSVS 603
Db 633 MSDRVKATQAEQLS-----NELATERSTAQNESARQ---QLERQNKEL 675
Qy 604 STVPPI TNNTEDEDRNTPKTFALNPKTPMTVTAPMQMAMTPSLANKVSATPVSIVYDK 663
Db 676 RSKLHEMGAIV--KSKPKFTIAALEAK-----IAQLEEQVEQEAEREQAATKSLKQDK 727
Qy 664 --PEVTLQEDIDYSEERRLA 682
Db 728 KLKEILLQVE-----DERQMA 743

RESULT 15

US-08-466-390-4
; Sequence 4, Application US/08466390
; Patent No. 568562
; GENERAL INFORMATION:
; APPLICANT: TOUTKATLY, GARY
; APPLICANT: TOUTKATLY, GARY
; TITLE OF INVENTION: NOVEL MALIGNANT CELL TYPE MARKERS OF THE
; TITLE OF INVENTION: INTERIOR NUCLEAR MATRIX
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,390
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: PITCHER ESQ. EDMUND R
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: MTP-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-390-4

Query Match 5.9%; Score 207; DB 1; Length 2101;
Best Local Similarity 21.1%; Pred. No. 5.4e-07;
Matches 123; Conservative 95; Mismatches 137; Indels 168; Gaps 24;
Qy 6 KQQLHQMSTTCDLSLLLELNVWDEVGEPDTRDRML-----LELEQECLEVYRRKVDQANR 61
Db 461 KQQLSSLTITDQSSISNLSQAKELEQASQAHGARLTAQVASTSE-LTTLNATIQDQDQ 519
Qy 62 SRAQLRKAIAEGEAEAGLACISAMGEPVHVHVSQKQLHGLREELNAIVPVLKEMKKKVE 121
Db 520 ELAGLQKQAEKQKQALQAT-----LQQEQASQGLRHQVLEQL---SSSLKQKE-- 564
Qy 122 RWNQFVHVIEQIKKISSE---IRPADVPFVKVVDQSDLRLKLDLTKDLESLOKESKD 178
Db 565 -----QQLKEVAEQEATRODHAQQLATAAEEREASLRERDAALKQLEALEKKA 615
Qy 179 RLKQVIEHLNSLSLCEVLGIDFKQTVYEVHPSLDEAGSKNLSNNTTIERLAAANRLRE 238
Db 616 KLEITLQQLQ-----VANEARDSAQTSTVTAQREKAELSR--- 650
Qy 239 MKIORMQKLODFASSMLELNLMDTPLEEQMFQ-----NITCNIAASEQITEPNTLSTD 294
Db 651 -KVEELQA-----CVETARQEQHEAQVAELEQLRSEQQKATEKERV AQ- 695
Qy 295 FLNVYSEVLRLEQLKASKMKDLVLKKALEEHRRRAHLVGBEGYAEBSIEAIEAGAI 354
Db 696 -----EKDQLQ-EQLQA--LKESLKVTKGSLSEKRA----- 725
Qy 355 DPSLVLEQIEAHATVKEBAFSPKDIIEKVERWQNAACEEAWLEDYKNDNRNAGR--- 411
Db 726 --ADALEEQQCISELKAET---RSLVEQHKKRERKELEEE-----RAGRKGL 767
Qy 412 -GAHLTLKRAEKARTLV-----NKIPGMVDVLTAKIAWKNRGEKEDFTYDGV 458
Db 768 EARLLQLGEAQAEFVLRRELAEAMAAQHTASECEQLVKEVAAMRD--GYEDSQOEEA 825
Qy 459 SLSSMLDEYMFVROEKEKQRDOKQLDQKAEQEAALYGSKPSPKPLS---TKKAP 515
Db 826 QYGMFQEQMLTLKE-ECEKARQ-----ELQEAKEKVAGIESHSELQISROQNKLA 876
Qy 516 RSHMGGANRLSLGGATMQPPKTDILHKSVSRAAKTTEETIGTL 558
Db 877 LHA-----NLARALQ-----VOEKEVRAQKLADLSTL 905

Search completed: November 20, 2004, 00:43:11
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 20, 2004, 00:37:56 ; Search time 85 Seconds

(without alignments)
2874.683 Million cell updates/sec

Title: US-10-619-685-2

Perfect score: 3496

Sequence: 1 MSSAVKQDLHQMSTTCDLSL.....DIDYFERRLAIFYLARQMV 690

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3496	100.0	690	15	US-10-619-685-2
2	2975.5	85.1	700	15	Sequence 2, Appl1
3	2939	84.1	592	17	Sequence 62819, A
4	1314.5	37.6	580	15	Sequence 253246,
5	1265.5	36.2	576	16	Sequence 268883,
6	1260	36.0	582	17	Sequence 131367,
7	1233.5	35.3	511	15	Sequence 223724,
8	1225	35.0	602	15	Sequence 49169, A
9	1213	34.7	630	16	Sequence 254332,
10	1187.5	34.0	598	17	Sequence 200324,
11	1182.5	33.8	589	16	Sequence 194860,
12	1181.5	33.8	559	16	Sequence 135276,
13	1141	32.6	660	17	Sequence 179742,
					Sequence 287895,

Sequence 106958,	14	1127.5	32.3	610	16	US-10-437-963-106958
Sequence 264132,	15	1101.5	31.5	611	17	US-10-425-115-264132
Sequence 354402,	16	1079	30.9	578	17	US-10-425-115-354402
Sequence 284613,	17	1077.5	30.8	953	17	US-10-425-115-284613
Sequence 64857, A	18	1074	30.7	603	15	US-10-425-114-64857
Sequence 261129,	19	1060.5	30.3	598	17	US-10-425-115-264129
Sequence 147340,	20	1042	29.8	553	16	US-10-437-963-147340
Sequence 312927,	21	960.5	27.5	1010	17	US-10-425-115-312927
Sequence 128829,	22	928	26.5	709	16	US-10-437-963-128829
Sequence 198413,	23	907.5	26.0	185	16	US-10-437-963-198413
Sequence 186153,	24	895.5	25.6	606	16	US-10-437-963-186153
Sequence 65219, A	25	874.5	25.0	554	15	US-10-425-114-65219
Sequence 134843,	26	873	25.0	176	16	US-10-437-963-134843
Sequence 17753,	27	870	24.9	538	16	US-10-437-963-17753
Sequence 281618,	28	799	22.9	234	15	US-10-424-599-281618
Sequence 183850,	29	709.5	20.3	147	16	US-10-437-963-183850
Sequence 42188, A	30	679.5	19.4	354	15	US-10-425-114-42188
Sequence 287893,	31	671.5	18.2	220	17	US-10-425-115-287893
Sequence 268884,	32	632	18.1	290	15	US-10-424-599-268884
Sequence 253791,	33	546.5	15.6	569	17	US-10-425-115-253791
Sequence 186497,	34	528.5	15.1	180	15	US-10-424-599-186497
Sequence 24267, A	35	489	14.0	315	16	US-10-767-701-24267
Sequence 56338, A	36	479.5	13.7	251	16	US-10-767-701-56338
Sequence 195661,	37	478	13.7	195	17	US-10-425-115-195661
Sequence 216368,	38	470.5	13.5	141	17	US-10-425-115-216368
Sequence 228892,	39	457	13.1	125	15	US-10-424-599-228892
Sequence 41154, A	40	430.5	12.3	325	16	US-10-767-701-41154
Sequence 50583, A	41	429.5	12.3	164	16	US-10-767-701-50583
Sequence 257687,	42	394	11.3	143	17	US-10-425-115-257687
Sequence 148583,	43	382.5	10.9	222	15	US-10-424-599-148583
Sequence 261176,	44	350	10.0	172	15	US-10-424-599-261176
Sequence 431, App	45	323.5	9.3	620	17	US-10-733-878-431

ALIGNMENTS

RESULT 1
US-10-619-685-2
; Sequence 2, Application US/10619685
; Publication No. US20040096875A1
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agrobiological Resources, Ministry of Agriculture,
; APPLICANT: Forestry and Fisheries
; APPLICANT: Japan Science and Technology Corporation
; TITLE OF INVENTION: A NOVEL GENE FOR CONTROLLING LEAF SHAPES
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/619,685
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: US/09/667,475D
; PRIOR FILING DATE: 2002-11-13
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-619-685-2

Query Match	100.0%;	Score	3496;	DB	15;	Length	690;
Best Local Similarity	100.0%;	Pred. No.	7.1e-216;				
Matches	690;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1
DB	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1	MSSAVKQDLHQMSTTCDLSLLELNVIWDEVEGPD	1
QY	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQKQLHGLREELNAIVPYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQKQLHGLREELNAIVPYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQKQLHGLREELNAIVPYLEMKKKV	61
DB	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQKQLHGLREELNAIVPYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQKQLHGLREELNAIVPYLEMKKKV	61	RSRAQLKAIAGEAEAGELAGICSAMGPPVHVRSQKQLHGLREELNAIVPYLEMKKKV	61
QY	121	ERNWQFVHTEQTKTSSSEIRPADFPVFKVPVQSDLSLRKLDELTKDLESLOKESDRL	121	ERNWQFVHTEQTKTSSSEIRPADFPVFKVPVQSDLSLRKLDELTKDLESLOKESDRL	121	ERNWQFVHTEQTKTSSSEIRPADFPVFKVPVQSDLSLRKLDELTKDLESLOKESDRL	121

```

Db 121 ERWNOFVHVIEQIKKISSEIRPADFVPFKVPVQSDLSLRKLDLTKDLESLOKEKSDRL 180
Qy 181 KOVIEHLNLSHLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 240
Db 181 KOVIEHLNLSHLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 240
Qy 241 IQRMQKLDQFASMSLELWNLMDTPLEEQMFQNTCNIAASEQITEPTNTLSTDFLNVE 300
Db 241 IQRMQKLDQFASMSLELWNLMDTPLEEQMFQNTCNIAASEQITEPTNTLSTDFLNVE 300
Qy 301 SEVLRLQKASMKDVLKKAELHRRRAHLVGEAGYAEFSEIAEAGADPSLVL 360
Db 301 SEVLRLQKASMKDVLKKAELHRRRAHLVGEAGYAEFSEIAEAGADPSLVL 360
Qy 361 EOIEAHIAATVKEAFSRKDIKLEKVERWQNAEEAWLEDYKNDNRYNAGRAHLTLKRA 420
Db 361 EOIEAHIAATVKEAFSRKDIKLEKVERWQNAEEAWLEDYKNDNRYNAGRAHLTLKRA 420
Qy 421 EKARTLVNKPIMGVVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKRR 480
Db 421 EKARTLVNKPIMGVVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKRR 480
Qy 481 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db 481 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Qy 541 LHSKSVRAAKTEETIGTISPSRRGLDIAGLPIKKLSFNASTLRETETPRKPPAQITPGN 600
Db 541 LHSKSVRAAKTEETIGTISPSRRGLDIAGLPIKKLSFNASTLRETETPRKPPAQITPGN 600
Qy 601 SVSSTPVRPITNNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 660
Db 601 SVSSTPVRPITNNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 660
Qy 661 YDKPEVTLOEDIDYSFEERRLAIVLARQV 690
Db 661 YDKPEVTLOEDIDYSFEERRLAIVLARQV 690

RESULT 2
US-10-425-114-62819
; Sequence 62819, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 62819
; LENGTH: 700
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3332-041-F11_FLI.pep
US-10-425-114-62819

```

```

Query Match 85.1%; Score 2975.5; DB 15; Length 700;
Best Local Similarity 85.1%; Pred. No. 1.9e-182;
Matches 588; Conservative 47; Mismatches 51; Indels 5; Gaps 5;

Qy 1 MSSAVKQDLQKQSTTCDLSLLELNVIVDEVEGPDPTTRDRMLLELQEGLEVYRRKVDQAN 60
Db 12 MSSAVKQDLQKQSTTCDLSLLELNVIVDEVEGPDPTTRDRMLLELQEGLEVYRRKVDQAN 71
Qy 61 RSRAQLRKAIBGEAEAGISAMGEPFVHVHRSQKQLHGLREELNVAIVPYLEEMKKKV 120

```

```

Db 72 RCRAQLRKAIBGEAEAGISAMGEPFVHVHRSQKQLHGLREELNVAIVPYLEEMKKKV 131
Qy 121 ERWNOFVHVIEQIKKISSEIRPADFVPFKVPVQSDLSLRKLDLTKDLESLOKEKSDRL 180
Db 132 ERWNOFVHVIEQIKKISSEIRPADFVPFKVPVQSDLSLRKLDLTKDLESLOKEKSDRL 191
Qy 181 KOVIEHLNLSHLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 240
Db 192 KOVIEHLNLSHLSCEVLGIDFKQTVYVHPSLDEAGSKVLSNTTIERLAAAANRLREMK 251
Qy 241 IQRMQKLDQFASMSLELWNLMDTPLEEQMFQNTCNIAASEQITEPTNTLSTDFLNVE 300
Db 252 IQRMQKLDQFASMSLELWNLMDTPLEEQMFQNTCNIAASEQITEPTNTLSTDFLNVE 311
Qy 301 SEVLRLQKASMKDVLKKAELHRRRAHLVGEAGYAEFSEIAEAGADPSLVL 360
Db 312 SEVLRLQKASMKDVLKKAELHRRRAHLVGEAGYAEFSEIAEAGADPSLVL 371
Qy 361 EOIEAHIAATVKEAFSRKDIKLEKVERWQNAEEAWLEDYKNDNRYNAGRAHLTLKRA 420
Db 372 EOIEAHIAATVKEAFSRKDIKLEKVERWQNAEEAWLEDYKNDNRYNAGRAHLTLKRA 431
Qy 421 EKARTLVNKPIMGVVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKRR 480
Db 432 EKARTLVNKPIMGVVDVLRTKIAAWKNERGKEDFTYDGVSLSSMLDEYMFVROEKEQKRR 490
Qy 481 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 540
Db 491 QRDQKLDQKAEQALYGSKPSKPLSTKKAAPRHSNGGANRRLSLGGATMQPPKTDI 550
Qy 541 LHSKSVRAAKTEETIGTISPSRRGLDIAGLPIKKLSFNASTLRETETPRKPPAQITPGN 600
Db 551 LHSKSVRAAKTEETIGTISPSRRGLDIAGLPIKKLSFNASTLRETETPRKPPAQITPGN 609
Qy 601 SVSSTPVRPITNNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 659
Db 610 SVSSTPVRPITNNTEDDENRTPKTTALNPKTPMTVTAPMQMAMTPSLANKVSPVSLV 668

RESULT 3
US-10-425-115-253246
; Sequence 253246, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 253246
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(690)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_162540C.1.pep
US-10-425-115-253246

```

```

Query Match 84.1%; Score 2939; DB 17; Length 690;
Best Local Similarity 84.4%; Pred. No. 4.1e-180;

```


	Matches	584;	Conservative	49;	Mismatches	53;	Indels	6;	Gaps	6;
Qy	1	MS	SAVKDQLHQMSTTCDLSLLLEINVIWDEVGPDPTTRDRMLLELSEOECLVEYRRKVDQAN	60						
				:		:		:		:
				:		:		:		:
Db	1	MS	SAKDQLQQNSTTVDSLLLELNMIWDEVGPDMDARDMLLEEQEMLDVYRRKVDQAN	60						
				:		:		:		:
				:		:		:		:
Qy	61	RRA	QRKAISAGEAELAGICSMAGEPPVHVQSQNKLHGLREELNAIVPYLEEMKKKV	120						
				:		:		:		:
				:		:		:		:
Db	61	RCA	QRQAIAADAELAELAGICSMAGEPPTHVRQSNQKLGHLREELNAIVPYLEEMRKKV	120						
				:		:		:		:
				:		:		:		:
Qy	121	EWN	NOPVHVIEOIKTSISSIRPADFVPFKVP--VDOSDSLRLKLDLTLDLESLOKEKSDR	179						
				:		:		:		:
				:		:		:		:
Db	121	ERW	DQVDVIEQIKKVAEIRSDFVFPIIPXVDOSDSLRLKLDLTLDLESLOKEKSDR	180						
				:		:		:		:
				:		:		:		:
Qy	180	LQV	IHLNSLSLCEVLGIDFKQTYYVHPHSIDEAESKNLSNTTIERLAAANLRBM	239						
				:		:		:		:
				:		:		:		:
Db	181	LQV	MEHLSTLSLCEVLGVDPFKQTVNEVHPSLGEADGSKNLSNCTIESLASAASRLREL	240						
				:		:		:		:
				:		:		:		:
Qy	240	KI	ORMQKLDQFASSMLELNLMDTPLEBQOMFONTCNTIAASEQEITENTILTSTDFLVY	299						
				:		:		:		:
				:		:		:		:
Db	241	KV	ORMQKLDQLASSMLELNLMDTPLEEBQOMFONTCNTIAASEHEITEPNTILSTDFLSYV	300						
				:		:		:		:
				:		:		:		:
Qy	300	ESE	VBLEQLKASKMDLVLKKAELIEHRRRAHLVGESGYAEFEFSIEAIEAGADPSLV	359						
				:		:		:		:
				:		:		:		:
Db	301	ESE	VBLEQLKASKMDLVLKKTTELDHRRRAHLVGESGYAAEFDEAIEAGADVPSLV	360						
				:		:		:		:
				:		:		:		:
Qy	360	LEO	IEAHIAIVKEAFSRKDILLEKVERWQACBEEAWLEDYNKDDNRYNAGRGAHLTLKR	419						
				:		:		:		:
				:		:		:		:
Db	361	LEO	IEAHIAIVKEAFSRKDILLEKVERWLNACEEAWLEDYNKDDNRYNAGRGAHLTLKR	420						
				:		:		:		:
				:		:		:		:
Qy	420	AEK	ARTLVNKIPGMVDVLRTKIIAAWKNERGKEDFTYDGVSLSSMLDEYMFVPRQEKQEK	479						
				:		:		:		:
				:		:		:		:
Db	421	AEK	ARILVNKIPLGDVLVTTKIWAWEARGKE-FTYDGVRLMSMLEEYIMVPRQELQKK	479						
				:		:		:		:
				:		:		:		:
Qy	480	ROR	DQKLDQLKAEQALYGSPSPSKPLSTKKAPRHSGGANRLSLGGATMQPKTD	539						
				:		:		:		:
				:		:		:		:
Db	480	ROR	DQKKIODQLKAEQALYGSPSPSKPSTKKAPRHSGGANRLSLGGATMQAPKTD	539						
				:		:		:		:
				:		:		:		:
Qy	5									

RESULT 4
 US-10-424-599-268883
 ; Sequence 268883, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)/B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 268883
 ; LENGTH: 582
 ; TYPE: PRT
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_84823C.1.pep
 US-10-424-599-268883

Query Match	37.6%	Score 1314.5	DB 15	Length 582
Best.Local Similarity	49.8%	Pred. No. 6.4e-76		
Matches	275	Conservative 100	Mismatches 140	Indels 39
Gaps				
Qy	15	TCDSLLLELNVIDEVGEPTTDRMLLELEQECLEVYRRKVDQANRSRAQLRAIAEGE	74	
Db	15	TCSSLLKKLQEIINDEVGESDEQRDKMLQLQECLELDVYKRVKEQAASRAQLQASDAK	74	
Qy	75	AEIAGICSAMGEPVHVHVRQSNQKHLGURELNAIVPYLEWKKKKVWRNQFVHVIEQIK	134	
Db	75	LELSTLLSALGKSFAGIPENTS-GTIKEQLAAIPVLEQLWQOKEERIKFSDVQSQIQ	133	
Qy	135	KISSEI-----RADFPVPKVPYDQDSLRLKLDLTKLDESLOKESDRLLKQVIE	185	
Db	134	QICCEIAGNUNLNDVSPA-----VDESLSLKKLDEYQSELEQLOKESERLHKVLE	185	
Qy	186	HLMSLSHLCSEVLGIDFKQTVYVHPHSUDEABG--SKNLSNTTIERLAAAANRLREMKTOR	243	
Db	186	FVSTVHDLCAVLGMDFFTATEVHPHSLNDSTGVQSKSIGSNDTLARLAKTVLTKEDKKOR	245	
Qy	244	MQKLODPASSMLELWNLMDTPEBQQFQNTCHIAASEQBITPENTLSTDFLNVYSESV	303	
Db	246	LHKLOELASQLIDLWNLDTHPEERRLFDHVTCHMSASVDVTPGALALDLIEQAEVEV	305	
Qy	304	LRLEQLKASRKMDLVLKKKAELEHRRRAHLVGEEVAAEFSEIAEAGADPSLVLEQI	363	
Db	306	ERLDQLKASRKKEIAFKQAELEIFARAHIEVDPDAAREKIMALIDSGNIEPTELLADM	365	
Qy	364	EAHIAIVKEBAFGRKOILEKVERMQNACEBAWLEDYKNDNRYNAGRGALHTLKRAEKA	423	
Db	366	DNQIAKAKEEALSRKDLDKVEKWSACEEBSWLEDYNRDRNRYNAGRGALHTLKRAEKA	425	
Qy	424	RTLWNKIPGMVDVLRTKIAAWKNRGKEDFTYDGVLSLSMLEDYMFVROEKEQKKRQD	483	
Db	426	RILWNKIPALVDTLTAVKTRAWEEHDHSM-SFTYDGVPLMLDDEYAMLHREEREEKKRMRD	484	
Qy	484	QKKLDOLKAEQAEALYGKPSPKPLSTTKA--PRHSMGGA-----NRLSLGGATMQPPK	537	
Db	485	QKHGEQRNTEQEIFGSRSPARPVSSSKSGGPR-ANGGANATPNRRLSLNA-----	536	
Qy	538	TDILHKSIVRAAKX	551	
Db	537	-----HONGNRSTSK	546	

RESULT 5
US-10-437-963-131367
; Sequence 131367, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131367
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_33439C.1.pgp
US-10-437-963-131367

US-10-424-599-268883
OTHER INFORMATION: Clone ID: PAT_MRT3847_84823C.1.pep
FEATURE:

```
Query Match          36.2%; Score 1265.5; DB 16; Length 576;
Best Local Similarity 44.3%; Pred. No. 8.5e-73;
Matches 269; Conservative 115; Mismatches 172; Indels 51; Gaps 11;

Qy 15 TCDSLLLELVINWDEGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLRAIEG 74
Db 15 TCDSLLLELVINWDEGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLRAIEG 74
Qy 6 TCGSLLOQLIWDVEGEDEDRKMLLQEQCLDVRKRVQDQATDSRDLIQAIDDSK 65
Db 6 TCGSLLOQLIWDVEGEDEDRKMLLQEQCLDVRKRVQDQATDSRDLIQAIDDSK 65
Qy 75 AELAGICAMGEPVHVROSQKLG-LREELNAIPVLEEMKKKVERWQFVHVIEQ 133
Db 75 AELAGICAMGEPVHVROSQKLG-LREELNAIPVLEEMKKKVERWQFVHVIEQ 133
Qy 66 IELARLLSALGEKA--IARTPEKTSGTIKQQLAAIAPTLEKLNKQKVERFVNVQSQI 123
Db 66 IELARLLSALGEKA--IARTPEKTSGTIKQQLAAIAPTLEKLNKQKVERFVNVQSQI 123
Qy 134 KKISSEIRPADFPVKV---PVDQSDLSLRKLDLTKDLESQKESDRLKQVIEHLNSL 190
Db 134 KKISSEIRPADFPVKV---PVDQSDLSLRKLDLTKDLESQKESDRLKQVIEHLNSL 190
Qy 124 DOI CCEIAGTTEVGEKVATPQVNEDDLTLERLEEFPSQLQELKEKSNRLEKVDYVSMI 183
Db 124 DOI CCEIAGTTEVGEKVATPQVNEDDLTLERLEEFPSQLQELKEKSNRLEKVDYVSMI 183
Qy 191 HSLCEVLGIDFKQTVYVHPSLDEAGS--KNLSNTTIERLAAAANRLREWKIQRMQKLG 248
Db 191 HSLCEVLGIDFKQTVYVHPSLDEAGS--KNLSNTTIERLAAAANRLREWKIQRMQKLG 248
Qy 184 HNLCTVLGMDFLSTVEVHPSLDDSIGDNCKSISNDTLKLDKTVATLNEKGRSLKLG 243
Db 184 HNLCTVLGMDFLSTVEVHPSLDDSIGDNCKSISNDTLKLDKTVATLNEKGRSLKLG 243
Qy 249 DFASSMLELWNLMDTPLEEQMFQNTICNIAASEQEITEPTNLSTDFLNVYVESEVLRLE 308
Db 249 DFASSMLELWNLMDTPLEEQMFQNTICNIAASEQEITEPTNLSTDFLNVYVESEVLRLE 308
Qy 244 ELAGQYDLWLDMDAPMOERSFHDVT CNRSASVDKVTAPGALADLLEQAEVEVQRUDQ 303
Db 244 ELAGQYDLWLDMDAPMOERSFHDVT CNRSASVDKVTAPGALADLLEQAEVEVQRUDQ 303
Qy 309 LKASKMDLVLLKKAELEHRRRAHLVGEEGYAEFEISIAEAGADPSLVLEQIEAHIA 368
Db 309 LKASKMDLVLLKKAELEHRRRAHLVGEEGYAEFEISIAEAGADPSLVLEQIEAHIA 368
Qy 304 LKYSKKEIAFKQTELEDIYAGAHWVDTAAAEKILALIEAGNIEPSELLADMESQIS 363
Db 304 LKYSKKEIAFKQTELEDIYAGAHWVDTAAAEKILALIEAGNIEPSELLADMESQIS 363
Qy 369 TVKEAFGRKOILEKVERQWACEEAEWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLVN 428
Db 369 TVKEAFGRKOILEKVERQWACEEAEWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLVN 428
Qy 364 KAKEEALSRKEILDKVERWMSACEEESWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLVN 423
Db 364 KAKEEALSRKEILDKVERWMSACEEESWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLVN 423
Qy 429 KIPGMVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKRRDOKK-L 487
Db 429 KIPGMVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKRRDOKK-L 487
Qy 424 KIPALVETLVAKTRAWESRGL-SPMYDGVPLLAMLDYVLMQRQESDKRMRQKRYI 482
Db 424 KIPALVETLVAKTRAWESRGL-SPMYDGVPLLAMLDYVLMQRQESDKRMRQKRYI 482
Qy 488 ODQLKAEQALYKSPSPKPLSTKAPRHSWGGA-----NRRSLGGATWQPPKTDIL 541
Db 488 ODQLKAEQALYKSPSPKPLSTKAPRHSWGGA-----NRRSLGGATWQPPKTDIL 541
Qy 483 EQQLNTDHEGFGSRVSNRPVSAKKVPGAKSNGSANGTTPNRRLSVSG----- 531
Db 483 EQQLNTDHEGFGSRVSNRPVSAKKVPGAKSNGSANGTTPNRRLSVSG----- 531
Qy 542 HSKSVRAAK--KTEIGTLSPSSRGLDIAGLPIKLSFNASTURETETPRKPAITPG 599
Db 542 HSKSVRAAK--KTEIGTLSPSSRGLDIAGLPIKLSFNASTURETETPRKPAITPG 599
Qy 532 HQNGRSGGKGRDRAKTSAPG-----NVAAKEDAS-----SHISGT 569
Db 532 HQNGRSGGKGRDRAKTSAPG-----NVAAKEDAS-----SHISGT 569
Qy 600 NSVSTP 606
Db 600 NSVSTP 606
Qy 570 DVPVSTP 576
Db 570 DVPVSTP 576

RESULT 6
US-10-425-115-223724
; Sequence 223724, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 223724
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_135624C.1.pep
US-10-425-115-223724

Query Match          36.0%; Score 1260; DB 17; Length 582;
```

```
Best Local Similarity 49.8%; Pred. No. 1.9e-72;
Matches 265; Conservative 99; Mismatches 152; Indels 16; Gaps 9;

Qy 14 TTCDLLELVINWDEGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLRAIEG 73
Db 14 TTCDLLELVINWDEGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLRAIEG 73
Qy 74 EAEIAGICAMGEPVHVROSQKLG-LREELNAIPVLEEMKKKVERWQFVHVIEQ 132
Db 74 EAEIAGICAMGEPVHVROSQKLG-LREELNAIPVLEEMKKKVERWQFVHVIEQ 132
Qy 74 KSEITRLLSALGES--SISGAPDKTSGTIKQLAAI SPSLETLLRKRKESRYKEFAHVQLQ 131
Db 74 KSEITRLLSALGES--SISGAPDKTSGTIKQLAAI SPSLETLLRKRKESRYKEFAHVQLQ 131
Qy 133 IKKISSIRPADFV--PFKVP-VQSDLSLRKLDLTKDLESQKESDRLKQVIEHLNS 189
Db 133 IKKISSIRPADFV--PFKVP-VQSDLSLRKLDLTKDLESQKESDRLKQVIEHLNS 189
Qy 132 IQILRDEIAGDLHIGEQLTEPHVNADLSVRRLNEVLSLQALQKEKSTRLSRILESVSS 191
Db 132 IQILRDEIAGDLHIGEQLTEPHVNADLSVRRLNEVLSLQALQKEKSTRLSRILESVSS 191
Qy 190 LHSICEVLGIDFKQTVYVHPSLDEAGS--SKNLSNTTIERLAAAANRLREWKIQRMQK 247
Db 190 LHSICEVLGIDFKQTVYVHPSLDEAGS--SKNLSNTTIERLAAAANRLREWKIQRMQK 247
Qy 192 VHDICSVLGMDFVGTIVTVHPSLDDSVGVQSKSISGETISKLRKVIIVLEBEKSRFAKI 251
Db 192 VHDICSVLGMDFVGTIVTVHPSLDDSVGVQSKSISGETISKLRKVIIVLEBEKSRFAKI 251
Qy 248 QDFASSMLELWNLMDTPLEEQMFQNTICNIAASEQEITEPTNLSTDFLNVYVESEVLRLE 307
Db 248 QDFASSMLELWNLMDTPLEEQMFQNTICNIAASEQEITEPTNLSTDFLNVYVESEVLRLE 307
Qy 252 QALASQSLDLWNLMDSPVEEQPFPHVTCIMSSLTDDVIVPGALADLVIOQAELEVERLD 311
Db 252 QALASQSLDLWNLMDSPVEEQPFPHVTCIMSSLTDDVIVPGALADLVIOQAELEVERLD 311
Qy 308 QLKASKMDLVLLKKAELEHRRRAHLVGEEGYAEFEISIAEAGADPSLVLEQIEAHIA 367
Db 308 QLKASKMDLVLLKKAELEHRRRAHLVGEEGYAEFEISIAEAGADPSLVLEQIEAHIA 367
Qy 312 QLKASKMDIAFKQTELEDIYARAHVAVDSSAARDRIMSI EESSPESELLADMENQI 371
Db 312 QLKASKMDIAFKQTELEDIYARAHVAVDSSAARDRIMSI EESSPESELLADMENQI 371
Qy 368 ATVKEEAFSRKOILEKVERQWACEEAEWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLV 427
Db 368 ATVKEEAFSRKOILEKVERQWACEEAEWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLV 427
Qy 372 LKANEESLSRKDILRVDNRMSACEEESWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLV 431
Db 372 LKANEESLSRKDILRVDNRMSACEEESWLEDYKDDNRYNAGRGGAHLTLKRAEKARTLV 431
Qy 428 NKIPGMVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKRRDOKK 487
Db 428 NKIPGMVDVLRTKIAAKWNERGKEDFTYDGVSLSSMLDEYMFVRQEKQEKRRDOKK 487
Qy 432 NKIPAI VDTLLAKTRAWESRGL-FTYDGVPLLAMLDYVLMQRQESDKRMRDOKM 490
Db 432 NKIPAI VDTLLAKTRAWESRGL-FTYDGVPLLAMLDYVLMQRQESDKRMRDOKM 490
Qy 488 ODQLKAEQALYKSPSPKPLST--KXAPRHSWGGA--RRLSL---GGAT 532
Db 488 ODQLKAEQALYKSPSPKPLST--KXAPRHSWGGA--RRLSL---GGAT 532
Qy 491 NEQLAAEQELFGSKSPARPOSSSRKRVAGARTNGGTPARRLSALQSGGT 542
Db 491 NEQLAAEQELFGSKSPARPOSSSRKRVAGARTNGGTPARRLSALQSGGT 542

RESULT 7
US-10-425-114-49169
; Sequence 49169, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49169
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMFLTEOSINTE071C04_FLI.pep
US-10-425-114-49169

Query Match          35.3%; Score 1233.5; DB 15; Length 511;
Best Local Similarity 50.5%; Pred. No. 8.2e-71;
Matches 252; Conservative 96; Mismatches 142; Indels 9; Gaps 6;

Qy 14 TTCDLLELVINWDEGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLRAIEG 73
Db 14 TTCDLLELVINWDEGEPTTDRMLLEQECLEVYRRKVDQANRSRAQLRAIEG 73
Qy 16 TTCGSLLOQLIWDVEGEDEDRKMLLQEQCLDVRKRVQDQATDSRDLIQAIDDSK 75
Db 16 TTCGSLLOQLIWDVEGEDEDRKMLLQEQCLDVRKRVQDQATDSRDLIQAIDDSK 75
```



```

126 FLVDVGRIGIKIISSEIRPANFDPFKVSVQDOSLSURKUEELRVELKSEKEGSRVKQWME 185
186 HJNSLSHSLCEVLGIDPKOTVYVEVHPSLDEAGSKNLSNTTIERAAAAANRLRMWKIORMQ 245
186 YLKTLSHLCVVLGVDFPKTISEIHSLSLDEAGSPNINNTTIEMLAWAIOLRBTIKORMQ 245
246 KLQDFASSMLELWNLMDTPLEEQQONFONITCNIAASBQEITEPTNLSTDFLNVSEVLRL 305
246 KLQDLASTLLELWNLMDTPFEEQAYQONITCNIAASBAELTEQNTLSIEFLNY----- 298
306 LEQLKASKMKDLVLKKAELEHRRRAHLVGEEGYAEFSEIAEAGAIQDPSVLQIEA 365
299 ----- 298
366 HIATVKEEAFSRKDILEKVERMQNACEEAWLEDYNDKDNRYNAGRGHAHLTLKRAEKART 425
299 -----DDNRYNAGRGHAHMLKRAEKARV 321
426 LVNKIPGMVDVL-----RTKIAAWKNRGEKEDFTYDGVSLSSMLDYMVFVQXEQEKKRQ 481
322 LVSKIIPGQTKPPPRSRLSQLSPRA-----IHRXPALSS-----HATHTRSR 367
482 RDQKLOQLOKABQALYGSFKSPKPLST--KKAPRHSMGGANRRLSLGGAT-MOPPKT 538
368 TNPAPHDPSPTQOTTL-----PVVSSPLASHRHRLPPHAFPSPRASYIIRGFTRIOPTKT 423
539 DILHKSVRKAAKTTEEICTLS-----PSSSRGLDI-----ACLPKILKLSFNA 580
424 T-----TSSSRSSSTTSFSMLAXLVSSWPCRCWFJLDMGNYLSCWTWAKVPGKGARVILPDD 480
581 STLREYETPRKPFPAQI-----TPGNSVSSTPVRPIINN-----TEDDENRTPKTFTALNPKTP 633
481 GGLRQVALPATA-AELMMDAPGCHFLADARAARVARLAALSABEELGNAVATPFMKEL 539
634 MVTAPMQMAWMTPSLAN-----KVSA-----TPVSLVYDKPEVTLQOEDID 673
540 GTPAPADMARLAAVATBEARSAKVAAVVAPPPPTFLOAEDDAPRLRDEMDV 593

```

RESULT 10

```

US-10-425-115-194860
; Sequence 194860, Application US/10425115
; Publication No. US2004021472A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 194860
; LENGTH: 598
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURES:
; OTHER INFORMATION: Clone ID: MRT4577_109
US-10-425-115-194860

```

Query Match	34.0%	Score 1187.5;	DB 17;	Length 598;
Best Local Similarity	42.6%	Pred.No. 9.le-68;		
Matches 259;	Conservative 118;	Mismatches 179;	Indels 51;	Gaps 10;

Qy	11	QMSTTCDLLELNVIWDEVGPEPTTRDRMLLEEQECLEVYRRKVDQANRSRAQLRKAI	70
:	:	: : :::	:::
Db	14	EAGTSCGALLRELQIIWAEVGESEGEKNALLDIERECELEVYRRKVDDANKRIVQLHQSV	73
:	:	: : :::	:::
Qy	71	AEGEAELAGTCSANGPEPPVHVROSQKLHGLREELNAIPVYLEEMKKKKVERWNQFHHI	130
:	:	: : :::	:::
Db	74	AAKEAEVASLMATIGEHLLYLK-KDGWPLKEQLATVAPVLSLKCKEERIKQFSDIR	132
:	:	: : :::	:::

```

Qy 131 EQIKKIS---SEIRPADVPFKVVDQSDLSLRKLDELTKDLESQKEKSDRLKQVIEHL 187
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 133 SQIEKIRFELSEYNDQGDPSLAAEEHDLSMRKLSNYQTQLRALOKDKSERLRKVLEVI 192
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 188 NSLHSLCEVLGIDFKQTVYEVHPSLDE--AEGSKNLSNTTITIERLAAAAANRLREMKIQMQ 245
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 193 NEVHSLCGVLGIDFGSTVHEVHPSLHQNGVEQSRIINSTLEGLASTTIYKLKAERKSRIH 252
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 246 KLDPFASMSLEMLWMDTPLEEQQOMFONITCNIAASEOEITEPNTLTSTDFLNYSVESVLR 305
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 253 KAREIMESLCOLWKLMDSPEEKQOFKQWSLLLPPEGITSPGVLSSETTEKMEFEVER 312
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 306 LEQLKASKMIDLULKKAELAEHRRAHLVGEEGVAEFSFEATAGAI DP SLVLEQIEA 365
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 313 LTRLKTSRLKEIVMKRAELEAICQNAHIEDPVSTAPEQT DAL IDSG LIDPS ELLANIES 372
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 366 HIATVKEEA FSRKDILKVERWQNACEEAWLEDYNKDDNRYNAGRGAHLTLKRAEKART 425
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 373 QILKAKEESLRKDIMRINRWIAACDEEAWLEENODPKRYASAGRGHAHLNKRAEKARI 432
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 426 LVNKIPGMVDVLRTKIAAWKNERKEDFTYDGVSLSSMLDEYMFVROEKEKKQRDOQK 485
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 433 LVRKIPSWDDLNRITFAWENARK-PFLYDGGR LISVLEYRLSRHQKESENRRYRDQK 491
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 486 KLDPQLKAEQALYGKSPSPKPLS-TYKA-----PRHSMMGGAN----RRLSLGGATMOPP 536
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 492 KJESSILLAEKAI PGSRPS PRKTSSLSRKANGYRPNGNTGLKPTPRRLSLGSATPE-- 549
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 537 KTDILHSKSVRAAKKTBEIGTLPSPSSRGL-----DIAGLPKULSFNASTL----- 583
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 584 RETETP 589
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||
Qy 590 SEPESP 595
Db   |||      : : |||      : : |||      : : |||      : : |||      : : |||

RESULT 11
US-10-437-963-135276
; Sequence 135276, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With Plant Growth
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ IDS: 204966
; SEQ ID NO 135276
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_36968C.1.pep
US-10-437-963-135276

```

RESULT 11

```

US-10-437-963_135276
; Sequence 135276, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 135276
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MFT4530_36968C.1 pep
US-10-437-963-135276

```

```

Query Match      33.8%; Score 1182.5; DB 16; Length 589;
Best Local Similarity 42.9%; Pred. No. 1.9e-67;
Matches 255; Conservative 133; Mismatches 176; Indels 31; Gaps 12;

QY  1  MSSAVRQDLQHMQMTTCDSLLELVNVLDVEGPEPTTRDRMLLELEQCELVYRKRKVDQAN  60
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db  1  MSALLRE-----TSCGSLTKQLQSVMDVEGSEEDRKVLFDDBCECLDVYRKRKVDQAT  54
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

Qy	61	RSRAQLRKAIAEGEAELAGISAMGEPPVHVHVRQSNQKLHGLRBEELNAINVPYLEEMKKVKV	120
Db	55	KSDLLLOALDYDKTELARLASALGEKSIDI-SPEKTARTIKQELTALAPTLEQLGCKKK	113
Qy	121	ERWNOFVHVIEQIKKISSEIRPADVPFKV---PVDQSDLSRLKLDLTOKLESLOKES	177
Db	114	ERIKELANIQSRIEIQIRGEIAGTLEMGQQVALPOINEDDUTVKLRREFQLOQLEBEKES	173
Qy	178	DLRKQVIEHLNLSHSLCEVLGIDFKQTVYVHPSLDEAEGS--KNLSNTTIERLAAANR	235
Db	174	RLREKVLHVGVHMDLQNVLGHDFRITIVQVHSSLDSDISGNEKHKNISNETLSKLDRTICT	233
Qy	236	IREMKIORMKQLQDFASSMLELNLMDPTLPEEQMFQNTICNTAASQEITETPTLSTDF	295
Db	234	LNEDKRLRLKQLQELATQYDLMDLMDTPTVEERSLFDHVS CNRTATV EEVWVP GALAVDV	293
Qy	296	LNMYSEVLRLBOLKASQKMDLVLKKAELBEHRRRAHLVGBEGYAEBSFIEAIEAGAD	355
Db	294	IDQAQTEVERLDQLKYSKMKTAFKQQAILEDIYASTHVVLDTAVAHEKIQALIESGME	353
Qy	356	PSLVLVEQTEAHTATVKEENAFSKOTLEKVEESQWQACEEAWLDEYKNDNRYNAGRAHL	415
Db	354	PSLELADMSQLTKAKEALGRKETLDKVERWISCSCEESWLDEYSDRDDNRYSNGRAHL	413
Qy	416	TLKRAEKARTLVNKIPGMVDVLRTKIAAWKNRCKEDFTYDGVSLSSMLDEYMFVRQKE	475
Db	414	NUKRAEKAILVSKLIPALVETLVAKTRAEENHGL-PPMVDGVSLLANLDEYVILKQERE	472
Qy	476	QEKRRORDQKLQDQ--LKAEEALYGSKPSKPLSTKK-APRHSMGGA-----NRRLSL	528
Db	473	EKKGRNREKQRTQELLNIDREGPGTRVFNVRTSAKVAGTKPNGGASNGTPSRRLST	532
Qy	529	GGATMOPPKTDILHKSQV--RAAKTTEETGTLSPPSSRGLDLAGLPIKULSFENAT	582
Db	533	G-----NOLNESKSTRSAGDKKKGA-SKNTATSLNEAA-PADKEAADST	577

RESULT 12
US-10-437-963-179742
; Sequence 179742, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION: Thomas J.
; APPLICANT: La Rosa, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179742
; LENGTH: 559
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77174C.1.pep
US-10-437-963-179742

	Query Match	33.8%	Score 1181.5;	DB 16;	Length 559;
	Best Local Similarity	46.6%;	Fred; No. 2e-67;		
	Matches 261;	Conservative 100;	Mismatches 166;	Indels 33;	Gaps 12;
QY	11 QMSTTCOSLLLELVNIVDEVGSPDTRDRMLLELLEQBCELEVYRKRVQDNRSRAQLKAI	70			
	: : : :	:	: :	:	:
DG	14 QMEPRREMLLGELGMEDQIGAEABDRREMILHALEEDCLNVYRKKVAVQVKGIIRAQLGREI	73			
	: : : :	:	: :	:	:
QY	71 AEGFAELAGICSAMGE--PVHYVRQSN-QKLHGLUREELNAITVPYLEEMKKXKVWRWNOFV	127			

Db		74	ADSAEVAALCATIGEPSTTVHACSLQSTGNLKEELSGSITPELEEMRRERRRKFS	133
Qy	8	128	HVIEIQIKKISSEIRPADFPVKVPVQDOSLSRLKLBELTKDLSLQKEKSDRLKQVIEHL	187
Db		134	EVTELINRIEQEMKPSK-QLHLTMDNSDLTIRRLBELRAYLODLOLEKDSRVKMTTELM	191
Qy		188	NSLHSLCEVLGIDFKOTVYVHPFSLDEAGSKNLSNTTIERLAAAANRLREMKIORMKL	247
Db		192	GSJHSSSLVLGMDFRETI--NLH---HDEG--DISDDAIARLVSIEIGRLREIKENRMOKL	244
Qy		248	QDPASSMLELWNLMDFLEBQQMFONITCNIAASEQSITEPTNLSTDFLNVSEVLRLE	307
Db		245	QDLATMLDNLMLDTPSEBQKRFQSVACNIAASEDEITERDALSMEFINNVEAEVVRLE	304
Qy		308	OLKASKMKDVLVKKKAELERHRRRAHLVGEGVAESEFIEAIEAGADPSLVLEQIEAH	367
Db		305	RLKECRMKDVLVKKYDELNRIERRAHVPVENEDAMMFDAISDA-KRSVLIERLEVQI	363
Qy		368	ATVKEEAFSRKDIKLEKVERWQNACEEAWLEDYKDDNRYNAGRGAHLTLKRAEAKARTLV	427
Db		364	SEAKDEEFSKDVLEKKEKQAALEESWLEEYNNRNNRYVNGKTHLVKBAEKARALV	423
Qy		428	NKIPGMVDVLRITKIAAMKNRGKEDFTYDGVSLSSMLDEYNFVROEKEKKGORDQKKL	487
Db		424	SKPMAMAEALITKVIAWEKERGAK-PFYDGDGLDMLDEYNNTRKEKEQERKRODQRM	482
Qy		488	QDLKAEQEAALYGSKPSPSKPLTKKAPRHSMGGANRRLSLGGATMQPPKTDILHKSQSV	547
Db		483	LQG-----GTGESF-----VVRPPPKNIKNVTRTLMSGTGTGCKA-----SASVS	524
Qy		548	AAKTEIEIGTLSPSSRRGLD	567
Db		525	SRPSTPSF-LKSPMSARRSD	543

RESULT 13

US-10-425-115-287895

; Sequence 287895, Application US/10425115

; Publication No. US20040214272A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants

; FILE REFERENCE: 38-21(53222)B

; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 287895

; LENGTH: 660

; TYPE: PRT

; ORGANISM: Zea mays

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(660)

; OTHER INFORMATION: unsure at all Xaa locations

; FEATURE:

; OTHER INFORMATION: Clone ID: MFT4577_25653C.1.pgp

US-10-425-115-287895

	Query Match	32.6%	Score 1141;	DB 17;	Length 660;
	Best Local Similarity	40.3%;	Pred. No. 1e-64;		
	Matches	257;	Conservative 115;	Mismatches 196;	Indels 70; Gaps 12;
Qy	15	TCDLLELNVIWDEVGEPDTRDRMLLEBQECLVYRRKVDQNRRAQLRKAIAGE	74		
		: : : : :			
Dd	74	TCGSLQLKLQFIWDEVGESIEDRDKLYQLQBCELDVYKRKVDATNSRDLLIQLDDSK	133		
		: : : : :			
Qy	75	AELAGICSAMGEPPVHVQRSNQKLHGRLNELNAIVPYLEEMKKKKVNRWNQPVHVEIQIK	134		
		: : : : :			

Db 134 IELARLLSALGEKAI-ARTPEKTTGTIKQOLAAIAPTLEQLTKQKNRKRREFVNVQSQID 192
Qy 135 KISSEIRPADFPFKV---PVDQSLSLRKLDELTKDLESLOKESDRKLKQVIEHLNSH 191
Db 193 QICGEIAGYVEGEQMTTQVNEDDTLTLEBEDFRSOLKDEKESXRLEKVLVDYVIVH 252
Qy 192 SLCEVLGIDFKQTVYVHPSLDEAGS---KNLSNTTIERIAAANRLREMKIORMOKLQD 249
Db 253 DLXCVLGMDFSLTVIEVHPSLDDSGDNCKSISNDLSKLDNTVATLNEDKKLRLSKQJE 312
Qy 250 FASSMELNMLMDTPLEBQOMFQNTICNIAAEEQETERTNTLSTDFLNVYVESEVLRLEOL 309
Db 313 LAGQLVDLMDLMDAPKEERMFPHVTCNRSASVDEVTAPGSLALDLIEQAEVEVQRLDOL 372
Qy 310 KASKMKDLVLKKAELERHRAHLNGEGYAEESIEAIEAGATDPSVLVLEQIEAHAT 369
Db 373 KYSMKEIAFKQNELEDIYAGAHIVIDTAAAHKEKIALIEAGNTEPSESLIADMOAQIAK 432
Qy 370 VKEEAFSRKDIIEKVERWQNAEEBAWLEDYNNKDDNRVYAGAHILTKRAEKARTLVNK 429
Db 433 AKKEALSRLDKLVKVERWSACEEESWLEDYNNKDDNRVYNSRGHNLNERAEAGEGSKK 492
Qy 430 -----IPGMVDVLRTKIAANKNERGKEDFTYDGVLSLMDLDEYMFVQEKQKORD 483
Db 493 DSAKTASPGNSAAATAPVGL--TATAKED-----AALLAMLDYVYMLRQEREERKKRMRE 545
Qy 484 QKK-LQDLKABOEALYKSPSPKPLSTKKAAPRHSWGGA-----NRRLSGATWOP 535
Db 546 QKRYVEQQLNTHDEGPFGRVFNRPASSKKAIGPKLNGSVNGTPPNRRUSISQ----- 601
Qy 536 PKTDILHKSVRAAKTEBIGTLPSSSRGLDIAPLPIKLSFNASTLRLETETPRKPPAQ 595
Db 602 -----QNGGGHGVRSKGKDKK-----DTAK-----T 623
Qy 596 ITPGNSVSTPVRPITNNVEDDENPTKFTALNPKTP 633
Db 624 ASFGNSAAATAPVGLTATAKEDA-ASQISVTDLVPESTP 660

RESULT 14

US-10-437-963-106958
; Sequence 106958, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106958
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11356C.1.pap
US-10-437-963-106958

Query Match 32.3%; Score 1127.5; DB 16; Length 610;
Best Local Similarity 41.4%; Pred. No. 6.6e-64;
Matches 260; Conservative 107; Mismatches 212; Indels 49; Gaps 14;

Qy 1 MSSAVKQDLQHWSTTCDSLLLELVNVDVEGPDPTTRDRMLLEOECEVYVRKVDQAN 60
Db 11 MDKAVRSSV-SFETPCGRLLELEQIWIETIGEREDKDRMFLETECNVYRRKVDSAN 69

Qy 61 RSRAQLRKAIAEGEAEAGLACISAMGE--PPVHVRSQSKLHGLREELNAIVPYLEEMKKK 118
Db 70 AERSQLRQSLMAXEAEALKVLVASIGEITPKPKVDEK-----QSLKEQLAKVTPLEDLSRK 125
Qy 119 KVEEWNQFVHVIEQIKKISSEI----RPADFPVPKVPVDQSLSLRKLDELTKDLESLOK 174
Db 126 KEERIKFSLVQSQIEKIKAKISDHNNQHDGPNVHSDKNDHDLSTRRLSDLOAERNLQK 185
Qy 175 EKSRLQVIEHLNLSHLCEVLGIDFKQTVYVHPSLDEA--EGSKNLSNTTIERIAA 232
Db 186 EKSDRLQKVITYVDEVHCLCSVLGMDFAKTVKDVHPSLHGANSSENSTNISDSTLEGLTET 245
Qy 233 ANRLREMKIORMOKLQDFASMLELWNLMDTPLEBQOMFQNTICNIAAASQEIETPTLS 292
Db 246 ILKLAERKTRVSKLQEIIVGKLHKLWNLMESTEQERRHFTRVAAVLGSSTEEITSSVLS 305
Qy 293 TDFLNVYSEVLRLEQLKASKMKDLVLKKAELERHRRRAHLNGEGYAEESIEAIEAG 352
Db 306 LETTQETEEVERLTQKASRMKELVLKKELEEDICSNAHMEPDMSTAPEKITALIDSG 365
Qy 353 AIDPSVLVLEQIEAHATVKEEAFSRKDIIEKVERWQNAEEBAWLEDYNNKDDNRVYAGRG 412
Db 366 LVDPCELLSSITETQIAKARBSLTKDIMEKVDRLWSACDEETWLEBYNQDSSRSYAGRG 425
Qy 413 AHLTLKBAEKARTLVNKGPMVDVLRTKIAANKNERGKEDFTYDGVLSLMDLDEYMFVRQ 472
Db 426 AHINLKAERKARILVQKIPSMIDNLIAKTFAWEDER-SVPFLYDGLARLVAILEEQKLRV 484
Qy 473 EKEEKKRQRDQKQLDQKAEALYKSPSPKPLS-TKKAPRHSWGGAANRRLSLGGA 531
Db 485 QKEEDKRRHRDQKQLQSLLLKEXELIFGSKSPKTSFNRTSSHPNG-----NGA 537
Qy 532 TMQPPKTDILHKSVRAAKTEBIGTLPSSSRG-----LDIAGLPIKLSFNAST 582
Db 538 GFMTVP-----PRVSAGSATPEL--LTPRSYSGRYNNYFKENRRLAAAPL-----NFST 585
Qy 583 LRETETPRKPPAQITPGNSVSTPVRPI 610
Db 586 VSK-EDSMSSSFASIS-----GSEPDSP 607

RESULT 15

US-10-425-115-264132
; Sequence 264132, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 264132
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_172503C.1.pap
US-10-425-115-264132

Query Match 31.5%; Score 1101.5; DB 17; Length 611;
Best Local Similarity 41.1%; Pred. No. 3.1e-62;
Matches 245; Conservative 113; Mismatches 207; Indels 31; Gaps 9;

Qy 1 MSSAVKQDLQHWSTTCDSLLLELVNVDVEGPDPTTRDRMLLEOECEVYVRKVDQAN 60
Db 11 MDKALRASV-SFTPTPCGALLRELEQIWIETIGEREDKDRMFLETECNVYRRKVDSAN 69

Search completed: November 20, 2004, 00:44:43
Job time : 89 secs

THIS PAGE LEFT BLANK